Unraveling the puzzles of spectroscopy-based non-invasive blood glucose detection

By

Ishan Barman

S.M., Mechanical Engineering

Massachusetts Institute of Technology, 2007

B.S., Mechanical Engineering

Indian Institute of Technology, Kharagpur 2005

SUBMITTED TO THE DEPARTMENT OF MECHANICAL ENGINEERING IN PARTIAL FULLFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN MECHANICAL ENGINEERING

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2011

Author .....
Department of Mechanical Engineering
March 31, 2011

Certified by .....
Robert J. Silbey
Class of 1942 Professor of Chemistry, Thesis Supervisor

Certified by .....
Peter T.C. So
Professor of Mechanical Engineering and Biological Engineering, Thesis Committee Chairman

Accepted by .....
David Hardt
Chairman, Department Committee on Graduate Students

To my parents, Asim and Mita Barman

Unraveling the puzzles of spectroscopy-based non-invasive blood glucose detection

by
Ishan Barman

Submitted to the Department of Mechanical Engineering
on May 5, 2011 in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in
Mechanical Engineering

ABSTRACT

Disorders of glucose homeostasis, including types 1 and 2 diabetes, represent a leading cause of morbidity and mortality worldwide. Diagnosis and therapeutic monitoring of diabetes requires direct measurement of blood glucose. Regardless of the clinical test performed, however, withdrawal of blood is currently required for measurement of blood glucose levels. Non-invasive measurement of blood glucose levels is highly desired, given the large number of diabetics who must undergo glucose testing several times each day. In this context, near-infrared (NIR) Raman spectroscopy has shown substantial promise by providing successful predictions of glucose at physiologically relevant concentrations *in vitro* and even in individual human volunteers at single sittings. Nevertheless, prospective application of a spectroscopic calibration model - over a larger population or over several sittings - has proven to be challenging.

This thesis investigates the optical and physiological challenges that impede calibration transfer by introducing non-analyte specific variances. Specifically, we present major advances in four research directions. First, the effects of sample-to-sample turbidity induced variations in quantitative spectroscopy are studied. To account for these variations, a novel method, based on the photon migration theory, is proposed. We demonstrate that the proposed method can extract intrinsic line shapes and intensity information from Raman spectra acquired in a turbid medium thereby improving quantitative predictions significantly. Second, we quantify the sensitivity of Raman calibration models to endogenous fluorescence and its temporal quenching. Application of shifted subtracted Raman spectroscopy is proposed to reduce the possibility of spurious models developed on the basis of chance correlation between the concentration dataset and quenched fluorescence levels. Third, we solve the problem of physiological lag between blood and interstitial fluid glucose levels, which creates inconsistencies in calibration, where blood glucose measurements are used as reference but the acquired spectra are indicative of ISF glucose levels. To overcome this problem, we introduce a mass transfer-based concentration correction scheme and demonstrate its effectiveness in clinical studies. Finally, we propose a new design for fabricating a handheld Raman glucose monitor by employing excitation and detection of wavelengths selected on the basis of their spectral information content.

Based on the advances in instrumentation and methodology outlined in this thesis, we anticipate that our current clinical studies will establish the viability of Raman spectroscopy for non-invasive blood glucose detection.

Thesis Supervisor: Robert J. Silbey
Title: Class of 1942 Professor of Chemistry

Thesis Committee Chairman: Peter T. C. So
Title: Professor of Mechanical Engineering
and Biological Engineering

ACKNOWLEDGEMENTS

The four years of my doctoral studies have been remarkably exciting, eventful and instructive. It has been a whirlwind emotional ride, but a wonderfully rewarding one. And the due credit for this must go to all the truly remarkable people I have met in the process.

First, I would like to acknowledge the unwavering support and insightful guidance of my advisor, Professor Michael Feld (late). His confidence in his group members was a joy to behold. His larger-than-life persona will be missed but all who knew him will carry his image in their hearts. I know I will. During my time in the Spectroscopy Laboratory, and especially after Michael's death, Dr. Ramachandra Rao Dasari has been a mentor to me. His constant support and steadfast enthusiasm for my work belies all bounds of rational thinking. As other members of our laboratory will testify, without his resourcefulness and creative inputs our projects would have not proceeded at half the pace. Through my interactions with Michael and Ramachandra, I have grown as a researcher and as a person.

I am also greatly indebted to all the members of my thesis committee. Professor Robert Silbey, of "Alberty and Silbey" fame (the Bible of physical chemistry), kindly agreed to supervise me after Michael's untimely demise. It is my great privilege to have worked under his guidance. Professor Peter So has been my savior on both the research and departmental fronts multiple times. His ability to analyze the details and look at the big picture simultaneously have helped shape my own thinking and, by extension, this thesis. Professor Steve Jacques has injected periodic doses of reality into our blood glucose research project and his suggestions on focusing on the anatomical aspects of the problem have benefited us greatly. Professor Mark Bathe has been very generous with his time and suggestions and his support on vital occasions have leased a fresh lease of life towards my final graduation.

I also owe a great deal of gratitude to my SM thesis advisors, Professor Nam Suh and Professor Sang-Gook Kim. I learnt a substantial amount of design theory and electrochemistry during my time in the Park Center for Complex Systems including a uniquely 'decoupled' way of solving a problem. Professor Sujoy Guha of IIT Kharagpur deserves a special mention here because of the way he carefully nurtured my research interests before I knew what research really meant.

I would like to take this opportunity to thank all my labmates, past and present, including Taesik Lee, Steve Bathurst, A.J. Schrauth, Malin Kjellberg, Gyunyoung Heo, Condon Lau, Jelena Mirkovic, Sasha McGee, Zoya Volynskaya, Obrad Scepanovic, Yongjin Sung, Dan Fu, Zahid Yaqoob and Niyom Lue. Collectively, they have reduced the trials and tribulations of life at MIT, otherwise known as "drinking from the firehose". Our group meetings, which have fluctuated from one to nine hours, have been a source of much academic inspiration and critical

analysis, intermixed with healthy doses of fun and games. Thanks are also in order for Luis Galindo (our research engineer par excellence), Zina Queen and Leslie Regan for their gracious help and skillful maneuvering at times of greatest uncertainty.

In my time as a doctoral student, I have interacted particularly closely with Chae-Ryon Kong and Jeon Woong Kang. I value our friendships as well as our working relationships which I hope will continue for many years to come. And finally two people in my time at the laboratory deserve special mention. Gajendra Pratap Singh and Narahara Chari Dingari have been at the core of almost everything I have accomplished here. They have spent hours and hours working alongside me on the many crazy ideas that have arisen from our brainstorming sessions. I am indebted to them for providing me with patient hearing at all times and, in all honesty, I do not know what form this thesis would have taken without their support and co-operation.

I also wish to thank all my friends who have provided me with the necessary diversions, especially Shankhadeep, Anirban and Shiladitya. My interactions with them have helped me maintain a generally positive outlook, regardless of the situation at hand.

A special token of appreciation needs to go out to Damayanti Halder. I can scarcely believe that we have known each other for the better part of a decade now! She has provided me with much-needed boluses of enthusiasm and encouragement and has been a constant 'rock' at all times.

Finally, it is a Herculean task to accurately pen down my immense gratitude for my parents. They have been supportive, beyond reason really, of my many adventures in research. For as far back as I can remember, they have shared with me their enthusiasm for life and their dedication towards improving society. This thesis is dedicated to them for all the support, guidance and love that they have showered on me over the years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Disorders of glucose homeostasis, including types 1 and 2 diabetes, as well as gestational diabetes, represent a leading cause of morbidity and mortality worldwide. In the United States, an estimated 2-5% of all pregnant women develop gestational diabetes, which can result in significant morbidity to the mother and the fetus if it goes undetected. At the same time, the incidence of diabetes in general has been increasing at a significant rate, affecting nearly 25.8 million Americans in 2010 (8.3% of the population), and it currently stands as the seventh leading cause of death in the United States [1]. Ominously, 7 million of the aforementioned 25.8 million are unaware that they have such an affliction. On a worldwide scale, the most recent statistics suggest that the number of diabetics is likely to rise from 171 million people in 2000 to 366 million in 2030 - which frighteningly maybe an underestimate of future diabetes prevalence [2]. Finally, diabetes also poses an enormous economic challenge, as evidenced by the $174 billion total estimated cost in the US alone [3]. This prohibitive bill includes $116 billion in excess medical expenditures and $58 billion in reduced national productivity.

Clearly, diabetes presents a significant challenge to the human civilization as a whole and must be tackled simultaneously on various fronts including diagnostic, therapeutic and social awareness (particularly in developing countries). Unfortunately, at the present time, diabetes - which is a chronic metabolic disorder characterized by high blood glucose levels - has no well-established cure but must be treated with regular insulin injections and other related medications to maintain tight glycemic control. Consequently, careful management of blood glucose levels

through frequent monitoring is imperative to maintain the patient's quality of life. Indeed, failure to adequately monitor and regulate glucose levels can lead to severe secondary health complications in multiple organs such as the retina, nerves, kidneys and the circulatory system [4].

Diagnosis and therapeutic monitoring of diabetes requires direct measurement of plasma (or blood) glucose. Measurement of the fasting glucose level is the preferred test for diagnosis of diabetes in children and non-pregnant adults, while the oral glucose tolerance test (OGTT) is the preferred method of diagnosis for gestational diabetes. In addition, patients with an established diagnosis of insulin-dependent diabetes (all type-I and many type-II) require frequent glucose measurements for therapeutic monitoring. Regardless of the clinical test performed, withdrawal of blood or interstitial fluid is currently required for measurement of blood glucose levels. Given the current landscape of diabetes monitoring, non-invasive measurement of blood or interstitial glucose levels is highly desired.

Further, application of a noninvasive device for glucose monitoring during OGTT is of great clinical interest. Such a development could lay the foundation for a predictive assay, referred hereafter as a spectroscopic glucose tolerance test (SGTT), and in principle should be easier to develop because of the less stringent accuracy requirements. Finally, a non-invasive glucose monitoring device would also be very useful in the development of new technologies such as an artificial pancreas to help diabetic persons automatically control their blood glucose level. An artificial pancreas provides the substitute endocrine functionality of a healthy pancreas and requires a closed-loop system that links glucose measurement with computer-driven insulin infusion, so that no input is required from the user [5].

Over the past two decades, significant advances have been made in developing instrumentation and methodologies for minimally invasive and non-invasive measurement of blood glucose [6], as detailed in Chapter-2. In particular, application of optical techniques, such as near-infrared (NIR) absorption spectroscopy, optical coherence tomography (OCT), photoacoustic spectroscopy and polarization spectroscopy, to this important diagnostic problem has been widely investigated [7-9]. Our laboratory has pioneered the use of NIR Raman spectroscopy for various biomedical applications [10-19], including measurement of blood analytes [20-26]. NIR Raman spectroscopy, which combines the substantial penetration depth of NIR light (within the so-called "diagnostic window") with the excellent chemical specificity of Raman spectroscopy, offers a promising solution for noninvasive detection of blood glucose and other analytes.

Specifically, for blood glucose detection, several investigators including our own laboratory have reported Raman-spectroscopy-based glucose predictions at physiologically relevant concentrations in serum [22], whole blood [25] and other *in vitro* samples, such as human eye aqueous humor [27, 28] (Chapter-3). Promising studies employing transcutaneous Raman spectroscopy in human volunteers have also been conducted [26, 29]. Despite these initial promising results in individual human subjects over limited periods of time, true prospective application of a glucose-specific calibration model (over a larger population or over prolonged periods of time) has proven to be challenging.

This thesis seeks to unravel some of the complex issues, from a physiological and spectroscopic standpoint, that continue to confound researchers. Further, we develop correction methodologies for several of these puzzling factors over the course of the thesis and demonstrate their efficacy in both *in vitro* and *in vivo* conditions. In addition, we present a design for a

miniaturized Raman sensor that can potentially open up our technology to the real end users of such a device, i.e. for self blood glucose monitoring.

We expect that the clinical feasibility studies currently in progress will substantially validate our improved instrumentation and methodologies and in combination will provide accurate prospective measurements in both normal and diabetic patients. We also believe that application of the new methodologies will not only help from a clinical accuracy standpoint but also provide greater insight into the fundamental physiology and physics of such measurements (for example, a deeper understanding of the physiological lag process and the changes induced by the onset of diabetes). Finally, the work presented in this thesis will also serve as an important resource for other researchers who are interested in performing similar blood analyte measurements and disease diagnosis using Raman spectroscopy *in vivo*.

1.1 Specific research questions

While it is evident that Raman spectroscopy can noninvasively provide information about blood glucose concentration (along with that of a host of other important analytes), the path to clinical translation is strewn with several technical challenges, which need to be overcome before further feasibility studies are implemented on human subjects. Indeed, a "black-box"/"brute force" approach to spectroscopy based blood glucose sensing is likely to create apparently functional models that are incapable of prospective prediction even in a single individual, due to the presence of spurious correlations. (More generally, the nearly $10 billion glucose sensing market has invited a host of ideas and proposals, several of which have turned into premature enterprises and almost universally failed to make a mark in the area. The trials and tribulations associated with the so-called "Holy Grail" of biophotonics is well-documented in the "The Pursuit of Noninvasive Glucose: Hunting the Deceitful Turkey" by John Smith [30].)

The goal of this thesis is to carefully examine these technical challenges and develop and validate methodologies to overcome these - unraveling one layer of complexity at a time. The specific research questions that are investigated in this work can be classified into the following categories:

Effect of turbidity induced spectral distortions and sampling volume variations

A major challenge in prospective application of transcutaneous Raman spectroscopy is the variation in the background (medium) optical properties. Absorption and scattering properties of skin vary from site to site, from subject to subject, and over time with aging and other physical alterations. Tissue optical properties are important because absorption and scattering determine the effective sampling volume for a given excitation wavelength and influence the magnitude of the collected Raman signal and, additionally, may have a wavelength-dependent component that causes local spectral distortions. A method to correct for tissue optical properties is therefore necessary for the success of a prospective study as they induce non-analyte specific variance in the developed calibration models.

Impact of non-linearity in the spectra-concentration relationship

Along with the variation in turbidity, fluctuations in several system and sample parameters (such as ambient temperature and temporal system drift) could cause the underlying assumption of linearity between the spectra and concentration datasets to fail. While in general weak non-linearities can be modeled by the conventional multivariate calibration methods by retaining larger number of factors in the reduced dimension data space than is necessitated by the chemical rank of the system, such an approach risks the inclusion of irrelevant sources of variance and noise in the calibration model. It is thus important to: (a) investigate the presence of non-

linearity, if any, and identify the possible sources of the same and (b) introduce a new modeling method that can account for such curved effects in a reproducible manner.

Fluorescence background and its variations over time

Fluorescence has long been a limiting factor for the extensive use of Raman spectroscopy as an analytical technique for a wide array of samples and applications. The fluorescence background present in the Raman spectrum acquired from a molecule often completely overwhelms the weaker Raman signal, excluding its use. This problem is often further exacerbated in biological samples, where autofluorescence from several other molecules in the sample, even those with low concentrations but relatively large fluorescence yields, can be difficult to remove. Further, methods of removing fluorescence can distort Raman spectral features. This can severely compromise multivariate calibration on a set of Raman spectra.

In addition to the broad fluorescence background and the accompanying noise (shot noise and detector noise), the prediction accuracy of a calibration model built for a set of Raman spectra can also be severely compromised by the quenching of the endogenous fluorophores over time due to sustained laser exposure. Such quenching can introduce spurious correlations between analyte concentrations and fluorescence levels that would preclude the possibility of prospective application. It is thus of great importance to develop a method to minimize the fluorescence and associated quenching while maintaining the spectral integrity of the Raman features.

Accurate reference concentration measurements

Another key factor that greatly affects the predictive performance is the accuracy of the reference measurements inputted into the calibration model as part of the training dataset. In nearly all

optical and spectroscopic studies, investigators use the blood glucose concentrations collected at definite time intervals as the reference concentrations in building the calibration model. However, because excitation at visible and NIR samples a tissue volume of approximately 1 mm^3 or less (except when such excitation is in the eye), a significant component of the *in vivo* Raman signal is likely to originate from the interstitial tissues rather than the intravascular compartment. Due to the slow diffusion kinetics, these two glucose values are typically not identical - especially during times of rapid change in glucose levels, which also (unfortunately) represents the time period over which calibration studies are performed. This lag time creates an inconsistency in spectroscopic calibration algorithms, which are based on reference blood glucose concentrations and the acquired tissue spectra. (Indeed, this experimental challenge is faced by virtually all non-invasive glucose sensing strategies.) Thus, the development of a correction strategy that can derive a suitable reference concentration - thereby ensuring consistency in the spectra-concentration relationship - is vital to the long-term success of the spectroscopic calibration model.

Prospects of spectroscopic system miniaturization for blood glucose detection

Finally, we will investigate possibly the single biggest hurdle in terms of translation of spectroscopic methods to the clinical setting, namely the complexity and lack of portability of the instrument due to its large (often prohibitive) spatial footprint. Indeed, even though Raman spectroscopy has come to be recognized as a powerful analytical tool for biomedical applications, there has not been any successful implementation for routine clinical diagnosis, especially as compared to other optical methods such as optical coherence tomography (OCT). Specifically, self blood glucose monitoring necessitates the development of a hand-held or a wearable device due to the suitability of such a device for home users. Such reduction in

footprint can also potentially drive down the costs associated with each unit. Clearly, an alternate system design is necessary to enable the systematic translation from the optical bench to the clinic and finally to the homes of individual patients.

1.2 Thesis outline

This dissertation is organized in the following manner:

In Chapter 2 we first explain the pathophysiology of diabetes mellitus as this understanding is crucial for motivating specific portions of the dissertation. We also briefly discuss other clinically relevant blood analytes, a number of which could be detected using the relevant methodologies outlined here. Finally, we describe the current clinical management of diabetes including existing diagnostic procedures, ranging from the gold standard tests to the minimally invasive ones and the proposed non-invasive approaches. A brief outline of therapeutic options is also provided to gain a deeper understanding of the accuracy and precision metrics necessary in the diagnostic methods.

Chapter 3 provides background information on quantitative biological Raman spectroscopy. We review the fundamental principles behind this inelastic scattering technique and the instrumentation necessary to acquire the experimental data. We discuss multivariate calibration modeling, an important element in extracting analyte concentrations from tissue Raman spectra. Previous accomplishments of this project are then summarized to establish the foundation of this dissertation work.

Chapter 4 introduces turbidity corrected Raman spectroscopy (TCRS) as a novel method to overcome the problem of turbidity induced sampling volume variations and spectral distortions. This method is based on the photon migration formulation and employs alternate acquisition of

diffuse reflectance and Raman spectra to compensate for the effects of turbidity in the acquired Raman spectra. The effectiveness of this method is validated through a series of experiments employing physical tissue models over a wide range of absorption and scattering properties.

Chapter 5 focuses on the presence of non-linearity in the spectra-concentration relationship. It is first demonstrated that the application of conventional linear models risks the incorporation of significant non-analyte specific variance and noise, especially in datasets acquired from human subjects *in vivo*. It is hypothesized that the curved effects could be partly attributed to tissue turbidity, especially when the effects of absorption are significant. Finally, a relatively new regression approach, known as support vector regression, is incorporated to demonstrate the potential of (reproducible) non-linear modeling for our specific application.

In Chapter 6, the effects of fluorescence and its variation over time are explored. In particular, it is demonstrated that sample photobleaching introduces spurious correlations in the calibration models for biological Raman spectroscopy, particularly in cases where temporal correlations exist within the concentration dataset. This has substantial implications for screening of diabetics, as any tolerance test based protocol would lead to an approximately monotonic rise in glucose levels over the measurement period due to the inadequate insulin response. Different fluorescence removal methodologies, both numerical and experimental, are examined as potential tools for avoiding the pitfalls associated with photobleaching correlated samples.

Chapter 7 presents a new dynamic concentration correction (DCC) scheme to account for the physiological lag between the glucose levels in the blood and interstitial fluid compartments, respectively. The proposed formalism allows the transformation of glucose in the concentration domain between the two compartments, ensuring consistency with the acquired spectra in the calibration model. Taking Raman spectroscopy as a specific example, it is shown that the

predicted glucose concentrations using DCC-based calibration model closely match the measured glucose concentrations, while those generated with the conventional calibration methods show significantly larger deviations from the measured values. Additionally, we develop an analytical formula for a previously unidentified source of limiting uncertainty arising in spectroscopic glucose monitoring from a lack of knowledge of glucose kinetics in prediction samples.

In Chapter 8, wavelength selection approaches are first employed in conjunction with support vector regression to ensure a robust and transferable calibration model by eliminating spurious pieces of information in the spectral dataset while allowing for the presence of curved effects. Subsequently, the impact of such a modeling approach on the development of a miniaturized Raman spectroscopic monitor for blood glucose is examined. Based on these principles, an alternate design embodiment, namely tunable excitation-single detection, is proposed. This eliminates the need for a conventional dispersive spectrograph-CCD combination by using serial acquisition of the sampled wavelengths. Such a design framework, coupled with the aforementioned wavelength selection, provides a substantive reduction in spatial footprint while maintaining acquisition times within required thresholds.

In Chapter 9, the major findings and conclusions of our investigations are summarized and the future directions for this work are discussed. We believe that further research along these specific directions, coupled with the work presented in this thesis, will finally enable the development of a clinically viable Raman spectroscopy-based blood glucose monitor.

1.3 References

(1) http://www.cdc.gov/diabetes/ (Accessed online on March 18, 2011)

(2) Wild, S.; Roglic, G.; Green, A.; Sicree, R.; King, H. *Diabetes Care* **2004**, 27,1047-1053.

(3) American Diabetes Association, *Diabetes Care* **2008**, 31, 596–615.

(4) Ross, S. A.; Gulve, E. A.; Wang, M. *Chemical Reviews*, **2004**, 104, 1255-1282.

(5) Sullivan, S. J.; Maki, T.; Borland, K. M.; Mahoney, M. D.; Solomon, B. A.; Muller, T. E.; Monaco, A. P.; Chick, W. L. *Science*, **1991**, 252, 718 - 721.

(6) Cunningham, D. D.; Stenken, J. A. *In vivo glucose sensing* **2009**, Wiley.

(7) Tuchin, V.V.; *Handbook of Optical Sensing of Glucose in Biological Fluids and Tissues* **2008**, CRC Press.

(8) Khalil, O. S. *Clinical Chemistry* **1999** 45, 165-177.

(9) Khalil, O. S. *Diabetics Technology & Therapeutics* **2004**, 660-697.

(10) Buschman, H.P.; Deinum, G.; Motz, J.T.; Fitzmaurice, M.; Kramer, J.R.; Laarse, A.; Bruschke, A.V.; Feld, M.S. *Cardiovascular Pathology* **2001**, 10, 69-82.

(11) Buschman, H.P.; Motz, J.T.; Deinum, G.; Romer, T.J.; Fitzmaurice, M.; Kramer, J.R.; van der Laarse A.; Bruschke, A.V.; Feld, M.S. *Cardiovascular Pathology* **2001**, 10, 59-68.

(12) Haka, A.S.; Shafer-Peltier, K.E.; Fitzmaurice, M.; Crowe, J.; Dasari, R.R.; Feld, M.S. *Cancer Research* **2002**, 62, 5375-5380.

(13) Shafer-Peltier, K.E.; Haka, A.S.; Fitzmaurice, M.A.; Crowe, J.; Myles, J.; Dasari, R.R.; Feld, M.S. *Journal of Raman Spectroscopy* **2002**, 33, 552-563.

(14) Motz, J.T.; Hunter, M.; Galindo, L.H.; Gardecki, J.A.; Kramer, J.R.; Dasari, R.R.; Feld, M. S. *Applied Optics* **2004**, 43, 542-554.

(15) Haka, A.S.; Shafer-Peltier, K.E.; Fitzmaurice, M.; Crowe, J.; Dasari, R.R.; Feld, M.S. *Proceedings National Academy of Sciences U S A* **2005**, 102, 12371-12376.

(16) Motz, J.T.; Gandhi, S.J.; Scepanovic, O.R.; Haka, A.S.; Kramer, J.R.; Dasari, R.R.; Feld, M.S. *Journal of Biomedical Optics* **2005**, 10, 031113.

(17) Haka, A. S.; Volynskaya, Z.; Gardecki, J. A.; Nazemi, J.; Lyons, J.; Hicks, D.; Fitzmaurice, M.; Dasari, R. R.; Crowe, J. P.; Feld, M.S. *Cancer Research* **2006**, 66, 3317-3322.

(18) Motz, J. T.; Fitzmaurice, M.; Miller, A.; Gandhi, S. J.; Haka, A. S.; Galindo, L. H.; Dasari, R. R.; Kramer, J. R.; Feld, M. S. *Journal of Biomedical Optics* **2006**, 11, 021003.

(19) Scepanovic, O. R.; Fitzmaurice, M.; Gardecki, J. A.; Angheloiu, G. O.; Awasthi, S.; Motz, J. T.; Kramer, J. R.; Dasari, R. R.; Feld, M. S. *Journal of Biomedical Optics* **2006**, 11, 021007.

(20) Berger, A. J.; Feld, M. S. *Applied Spectroscopy* **1997**, 51, 725-732.

(21) Berger, A. J.; Koo, T. W.; Itzkan, I.; Feld, M. S. *Analytical Chemistry* **1998**, 70, 623-627.

(22) Berger, A. J.; Koo, T. W.; Itzkan, I.; Horowitz, G.; Feld, M. S. *Applied Optics* **1999**, 38, 2916-2926.

(23) Koo, T. W.; Berger, A. J.; Itzkan, I.; Horowitz, G.; Feld, M. S. *Diabetes Technology & Therapeutics* **1999**, 1, 153-157.

(24) Hanlon, E. B.; Manoharan, R.; Koo, T. W.; Shafer, K.; Motz, J.; Fitzmaurice, M.; Kramer, J. R.; Itzkan, I.; Dasari, R. R.; Feld, M. S. *Physics in Medicine and Biology* **2000**, 45, R1-59.

(25) Enejder, A. M. K.; Koo, T. W.; Oh, J.; Hunter, M.; Sasic, S.; Feld, M. S.; Horowitz, G. L. *Optics Letters* **2002**, 27, 2004-2006.

(26) Enejder, A. M. K.; Scecina, T. G.; Oh, J.; Hunter, M.; Shih, W. C.; Sasic, S.; Horowitz, G.; Feld, M. S. *Journal of Biomedical Optics* **2005**, 10, 031114.

(27) Lambert, J. L.; Morookian, J. M.; Sirk, S. J.; Borchert, M. S. *Journal of Raman Spectroscopy* **2002**, 33, 524-529.

(28) Lambert, J. L.; Pelletier, C. C.; Borchert, M. *Journal of Biomedical Optics* **2005**, 10, 1-8.

(29) Chaiken, J.; Finney, W.; Knudson, P. E.; Weinstock, R. S.; Khan, M.; Bussjager, R. J.; Hagrman, D.; Hagrman, P.; Zhao, Y.W.; Peterson, C. M.; Peterson, K. *Journal of Biomedical Optics* **2005**, 10, 031111.

(30) Smith, J. L. *The Pursuit of Noninvasive Glucose: "Hunting the Deceitful Turkey"*, **2006**.

CHAPTER 2

BLOOD GLUCOSE AND DIABETES

Blood analytes provide valuable information for the diagnosis of many diseases and related health conditions. In fact, blood is inarguably the most important body fluid from a diagnostic perspective and the most widely tested in clinical chemistry because of the presence of a vast number of components whose concentration levels are tightly controlled under normal circumstances. Thus any deviation from these well established ranges can be immediately correlated with an abnormality in body function. Furthermore, blood tests are commonly used to evaluate effectiveness of drugs that have been delivered in response to a specific change of a physiological or/and biochemical state. Indeed, except in hematology, most of the routine 'blood' tests are used to assess the dissolved constituents in the plasma or serum, i.e. the liquid component of blood without the constituent cells (plasma) and often minus the clotting factors (serum) as well.

Typically, detection of a blood analyte would involve the withdrawal of blood from the test subject (often through venipuncture) and centrifugation to remove the cells before addition of specific chemical reagents. While these techniques are very accurate and robust, they necessitate transport of the samples to a well-controlled testing facility and often multi-step processing before the concentration numbers become available to the attending physician. Evidently, non-invasive point-of-care testing of the clinically relevant blood analytes, potentially using optical techniques, is desirable.

In particular, non-invasive glucose detection has been extensively investigated as diagnosis and therapeutic monitoring of diabetes necessitates direct measurement of blood glucose. Furthermore, the frequency of measurements necessary for glucose monitoring and the scale of diabetes in the global healthcare context provide greater clinical significance to this particular problem in relation to the noninvasive measurement of other important analytes such as urea and creatinine. Indeed, the problem of non-invasive glucose detection is often referred to as the "Holy Grail" of biophotonics.

This chapter presents background information on glucose and diabetes mellitus. The origin of different kinds of diabetes is discussed along with the sensing requirements specific to each type. Subsequently, we outline the requirements for a clinically viable glucose sensor and review the existing as well as proposed technologies for glucose detection. The goal of this chapter is to introduce the reader to the major glucose sensing techniques, which are selected based on their novelty, potential impact to the field and presence of preliminary results. For a more exhaustive coverage of these sensing techniques, the reader is referred to a series of excellent periodic reviews [1, 2] and compilations [3, 4].

2.1 Glucose and *diabetes mellitus*

Glucose represents the omnipresent fuel in biology and is used by cells as a source of energy as well as a metabolic intermediate. It is a monosaccharide and primarily exists as a closed-chain derivative of D-glucose, namely pyranose, in aqueous solution (Fig. 2.1). The mirror image of D-glucose, L-glucose, is rarely found in nature and cannot be used by mammalian cells [5]. The glucopyranose structure can be further designated as α- or β-, depending on the relative configuration of the hydroxyl group attached to C-1 and the $-CH_2OH$ group at C-5 in the Haworth projection. Under room temperature conditions, the cyclic isomers of glucose (aq.

solution) undergo muta-rotation over a few hours resulting in a consistent ratio of α:β 36:64. The Raman spectrum of the anomeric balanced D-glucose is shown in Fig. 3.4.

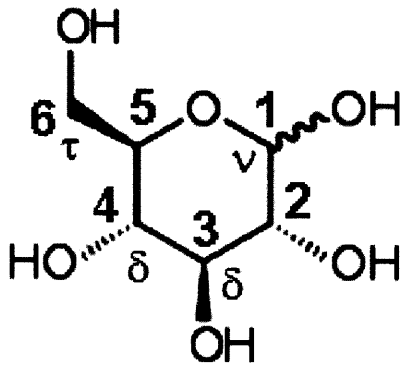


Fig. 2.1 Structure of D-Glucose. (The nomenclature of this figure is critical to the understanding of the corresponding vibrational spectra acquired from the aqueous solution of glucose, Fig. 3.4)

Glucose metabolism in the human body involves a complex interaction chain of a host of biological substances with the ultimate result that a reasonably constant blood glucose level is maintained at all times. In general, the instantaneous plasma glucose concentration, which can be viewed as a balance between the glucose entering the circulation and that being removed from the circulation, is closely regulated in the range of 45 to 180 mg/dL (2.5 to 10.0 mM) [6]. Glucose concentrations higher than the normal range are referred to as hyperglycemia and glucose concentrations lower than the normal range are referred to as hypoglycemia. (It is worth mentioning that a subtle difference exists between blood glucose and plasma glucose concentration. Typically, a multiplication factor of 1.15 is applied to the blood glucose level to convert to the corresponding serum/plasma level. In this thesis, we will hereafter refer to this measure as blood glucose concentration for the sake of consistency.)

Glucose, which enters the blood circulation, stems primarily from three sources, namely the intestinal absorption after a meal, glycogenolysis and gluconeogenesis. The latter two processes, often characterized as endogenous glucose production mechanisms, are controlled in part by the catabolic hormone glucagon (which opposes the effects of the primary anabolic hormone, insulin). On other hand, glucose removal from the circulation can be attributed to its consumption and use by the cells in the peripheral tissues. Terminal oxidation, glycolysis and storage as glycogen in the muscle and liver form the primary pathways for glucose metabolism.

In several of these aforementioned pathways, insulin plays a vital role through its complex chain of actions in various organs in the body. First, insulin facilitates the entry of glucose into muscle, adipose and several other tissues by binding to specific receptors (e.g. GLUT4, an insulin-regulated glucose transporter found in adipose tissue and striated muscle [7]). Second, insulin stimulates the liver to store glucose in the form of glycogen via the process of glycogenesis. Finally, it inhibits the glucagon secretion from pancreatic α-cells thereby terminating the endogenous production of glucose in the liver. The insulin action itself - both in terms of its synthesis in the β-cells of the pancreas and its release - is closely mediated by the concentration of glucose in circulation. For example, while insulin is not released if the blood glucose concentration falls below 3.3 mM, significantly higher levels (8 mM or more) may be subjected to a long-term release of insulin (first through the rapid release of preformed insulin and subsequently by the release of newly synthesized insulin).

Evidently, the chain of these antagonistic hormone-mediated reactions is predicated on the glucose concentration levels in the body, which fluctuate substantially over the course of a day primarily as a function of meal intake. Typically, the initial meal intake is presented as a sharp spike in the glucose level. Subsequently, once the insulin action is initiated, the glucose levels

start falling, albeit at a slower rate than during the initial rise phase. Importantly, this means that the blood glucose concentration measured from a human subject must be carefully interpreted in terms of the time delay between a meal and the particular measurement. Other factors, notably age and gender of the subject, may also affect the measured blood glucose value.

Given the necessity of maintenance of closely regulated glucose levels as a critical part of the overall metabolic homeostasis, violations of glucose metabolism can result in a wide variety of severe physiological repercussions. For example, the lack of insulin action can result in the breakdown of fats resulting in the production of toxic ketones, even if high levels of glucose exist in the bloodstream. The accumulation of ketones may result in life-threatening complications, referred to as diabetic ketoacidosis, which can only be alleviated by additional administration of insulin.

Nevertheless, a large majority of the complications can be traced, in one form or another, to *diabetes mellitus*, the most widespread and well-known disorder of glucose homeostasis. Diabetes, as it is commonly called, is a direct consequence of impaired insulin secretion and varying degrees of peripheral insulin resistance. This leads to a state of persistent hyperglycemia. There are three main categories of diabetes: Type 1 (Insulin Dependent Diabetes Mellitus, IDDM), Type 2 (Non Insulin Dependent Diabetes Mellitus, NIDDM) and gestational diabetes.

Type 1 diabetes is characterized by the loss of the insulin-producing β-cells of the islets of Langerhans in the pancreas, resulting in absolute insulin deficiency. The bulk of these cases are caused by the autoimmune-mediated destruction of the β-cells [8]. Since no insulin is produced in an endogenous manner, Type 1 diabetics must be administered exogenous insulin, usually through periodic injections. Type 1 diabetics make up approximately 10% of the total diabetic population but need a significantly greater share of the diagnostic and therapeutic resources

because of the greater frequency of necessary measurements and associated injections. The pathophysiology of Type 2 diabetes includes impairments in both insulin action and secretion [9-11]. In insulin-resistant conditions, the rate of clearance of glucose in response to insulin action reduces. Currently, there is no established cure for both Type 1 and Type 2 diabetes, although pancreas transplants have been tried with limited success in Type 1 diabetics and Type 2 may be controlled with medications.

The third category of diabetes, gestational diabetes, occurs in pregnant women, who have not been affected by any form of diabetes previously, but display high blood glucose levels during the period of pregnancy. An estimated 2-5% of all pregnant women develop gestational diabetes, which can result in significant morbidity to the mother and the fetus if it goes undetected. In principle, gestational diabetes is fully treatable and usually resolves after delivery. Despite its fully treatable status, approximately 20%–50% of the affected women develop Type 2 diabetes later in life.

Due to the chronic nature of both Type 1 and Type 2 diabetes, multiple complications may arise over the course of time. These complications are primarily vascular in nature affecting both the smaller capillaries and the larger blood vessels. Among vascular complications of diabetes, the most serious ones include diabetic retinopathy and cataract, diabetic nephropathy, neuropathy and cardiovascular disease. Diabetes may also cause acute complications, such as the previously stated diabetic ketoacidosis and hypoglycemia. Indeed, hypoglycemic cases most commonly occur as an acute complication of (excessive) treatment of diabetes with insulin or oral medications. Any combination of the acute complications may further result in diabetic coma. Ominously, an estimated 2 to 15 percent of diabetics will suffer from at least one episode of diabetic coma in their lifetimes as a result of severe hypoglycemia.

In the context of diabetes, the consensus opinion in the medical community is that the most serious complications can be substantively reduced by frequent measurements and tight control of blood glucose levels [12]. However, excessively tight control may lead to an increased risk of hypoglycemia which can also have serious consequences as discussed above. In other words, frequent monitoring - potentially even a continuous one - is necessary to design an appropriate control scheme for administration of insulin, especially for Type 1 diabetics. Additionally, people with Type 2 diabetes are also recommended to monitor their blood glucose levels at least once in a day. In fact, numerous studies and reports contend that more frequent measurements are actually necessary. To illustrate this more clearly, let us consider the results of a recent clinical study where the diabetic subject performed on average nine blood glucose measurements per day [13]. Despite this prohibitively large number of measurements, the diabetic subjects still spent 4.8 hours per day in hyperglycemia and 2.1 hours per day in hypoglycemia on average.

Moreover, as detailed by Wickramasinghe and co-workers [14], the accurate estimation of blood glucose concentration is critical in the management of preterm infants during critical care. Preterm infants suffer from high risk of hypoglycemia because of low glycogen reserves compounded with feeding difficulties. In this group of patients, frequent intermittent blood glucose estimations are performed using a heel prick - a painful and inconvenient method considering the patient status.

Based on the above discussion, it is amply clear that a clinically accurate non-invasive continuous blood glucose monitoring system is highly desirable. However, the state-of-the-art in glucose sensing only affords some of these features. In the next section, we outline the salient characteristics of the systems that are currently employed for glucose detection, both in clinical laboratories as well as in the home environment.

2.2 State-of-the-art in glucose sensing

Over the years, blood glucose monitoring has shown a discernible shift away from clinical laboratory measurements towards point-of-care testing. Nevertheless, before focusing on the commercially available glucose meters (also known as glucose biosensors), it is important to have a quick look at the different laboratory and functional tests performed in clinical chemistry to gain a broader understanding of the field of glucose sensing.

In order to assess the functional status of the blood glucose regulation system, multiple laboratory and functional examinations have been developed [12, 15]. These tests can be primarily classified into two categories. The first one involves the direct measurement of any sample constituent concentration in the relevant biological tissue, fluid or similar compartment of interest. Specifically, glucose can be measured in the blood, plasma, serum, urine, cerebrospinal fluid etc. for investigation of particular functions. Among the standard testing procedures, the fasting blood sugar test (FBS), urine glucose test and two-hr postprandial blood sugar test (2-h PPBS) would fall under this category. An interesting advance in this area is the introduction of the glycated hemoglobin (HbA1c) test. HbA1c is formed by the non-enzymatic glycosylation of hemoglobin exposed to blood glucose and therefore has a strong correlation with the average glucose concentrations in the bloodstream in the preceding three month period (life span of the erythrocytes). The 2010 American Diabetes Association Standards of Medical Care in Diabetes has incorporated the HbA1c performance as an additional criterion for the diagnosis of diabetes [16] but this idea has been not met with universal acceptance yet [17].

The other group of tests comprises the functional tests, which investigate the metabolic interactions within components of blood glucose regulation system. The central idea of these functional tests is to (externally) change the concentration of a specific constituent and

concomitantly monitor the response of the system to this change. This class of tests includes the glucose loading tests (also known as glucose tolerance tests and glucose challenge tests) and insulin or glucose clamping studies. These tests can either be performed orally or intravenously, depending on the type of change and response sought. The functional tests are widely used for screening gestational diabetes and remain the most viable protocol for calibrating new sensing technologies such as non-invasive optical sensors.

As far as glucose measurement in serum/plasma samples in the clinical laboratory is concerned, two primary methods have been employed (although several different versions of these two major techniques have been used at one point of time or another). The initial method of choice was based on the nonspecific reducing property of glucose in a reaction with an indicator substance that changes color when reduced. This method has largely been discontinued because other blood compounds such as urea also have reducing properties and the technique is thus susceptible to producing erroneous readings. More recently, the preferred method has been to exploit glucose-specific enzyme reactions, which afford significant specificity. For example, glucose can be detected by its reaction with glucose oxidase (enzyme), with the generation of gluconic acid and hydrogen peroxide. The enzyme is reoxidized with an exogenous mediator reagent, such as a ferrocene derivative or Os(III). The mediator in turn is reoxidized by reaction at the electrode, which generates an electrical current proportional to the amount of blood glucose that has reacted with the enzyme. Alternately, oxygen present in the sample can react with the reduced enzyme forming hydrogen peroxide. In such cases, the oxidation of hydrogen peroxide is monitored (in the same way as the re-oxidation of the mediator compound) by the resultant generation of an electric current proportional to the glucose reactant. This amperometric detection method is precise, linear, and free from important interferents. In particular, the YSI

(formerly Yellow Spring Instruments) blood glucose analyzers, which were launched in 1975 and were based on the amperometric detection of hydrogen peroxide, have become the standard for clinical diagnostic work in hospitals and clinical chemistry laboratories worldwide. Remarkably, the construction of these analyzers has barely changed in more than three decades!

As briefly mentioned earlier, the focus of glucose sensing has increasingly shifted towards point-of-care testing, i.e. from the clinical laboratory to the patient's home. In comparison with conventional methods of glucose testing and particularly the unwieldy blood glucose analyzers, present day blood glucose meters provide a rapid and convenient method of determining glucose concentrations. Indeed the blood glucose meters have been popularized to such an extent that at the time of this writing they occupy nearly 85% of the world market for *all* biosensors [18].

Blood glucose monitoring by such a meter involves pricking the finger with a lancet, extracting a drop of blood and depositing on a single-use test strip provided with the meter, and finally placing the strip into a meter for determination of the glucose level. While different methods of transduction including light reflectance based detection, electrochemical methods dominate the field of blood glucose meters due to their sensitivity, reproducibility and ease of fabrication. Fundamentally, the electrochemical detection scheme for these meters is not too dissimilar to the enzyme and mediator-based amperometric detection technique discussed above for the clinical laboratory testing. The typical enzymes used for the glucose meters are glucose oxidase and glucose dehydrogenase. The latter provides greater sensitivity over the former but is more vulnerable to interfering reactions with other biomolecules. The major difference between the clinical analyzers and the home blood glucose meters is that the latter are produced at a fraction of the size and cost of the former. The home blood glucose meters also employ only a

tiny volume of capillary blood, of the order of 0.3 to 1 μL (as opposed to a few hundred microliters required for the YSI analyzer).

An important point that bears discussion is the accuracy of these devices. According to International Organization for Standardization (ISO) 15197 guidelines, the prediction errors (in relation to a laboratory standard) for any meter device should be less than 15 mg/dL for any true glucose concentration below 75 mg/dL and should be less than 20% for any value higher than 75 mg/dL. A meter system is considered to meet the accuracy standards if 95% of pairs satisfy these criteria, separately for each range. In general, all the blood glucose meters on the market meet these accuracy standards, failing which Food and Drug Administration (FDA) approval would not have been granted. Nevertheless, it is noteworthy that several factors can potentially affect the accuracy of a specific test including the ambient temperature, size and quality of blood sample and high levels of interferents (e.g. acetaminophen and ascorbic acid) in the sample.

Currently, one can find more than 40 commercialized meter devices for detecting blood glucose levels. However, four multinational companies, namely Roche Diagnostics (*Accu-Chek* series), LifeScan (*One Touch* series), Abbott (*FreeStyle* and *MediSense* series) and Bayer (*Ascenia* series), dominate the market space. The widespread popularity of these devices and the enormous market size has prompted a significant portion of the current research efforts to be directed towards optimization of the existing technology. The recent advances in this area mainly concern the reduction of the manual operating steps in conventional lancet-based systems (lancet and lancing device), with the ultimate goal of obtaining an integrated device.

While the home blood glucose meters remain the current state-of-the-art, they are not by any means ideal. Indeed the market size and the huge demand created by an ever increasing diabetic population have encouraged researchers to develop novel painless approaches. Clearly, blood

glucose monitoring is an application area where opportunities for new technology abound. In the following sections, we will explore some of the innovative approaches which have the potential of changing the landscape of glucose monitoring in the coming years.

2.3 Minimally invasive glucose sensing techniques

Minimally invasive blood glucose monitoring techniques typically entail the insertion of an indwelling glucose monitor under the skin to measure the interstitial fluid glucose concentration. The electrochemical continuous glucose monitoring systems (including reverse iontophoresis-based sensors) as well as the optical sensors based on fluorescent probes and functionalized carbon nanotubes fall under this category. It is worth emphasizing that the technologies discussed below only provide a representative set of minimally invasive approaches under investigation; the reader may find it instructive to survey other sensing methods presented in the literature [19].

2.3.1 Electrochemical techniques

Perhaps the most attractive technique for invasive glucose monitoring is the electrochemical transducer, which may be implanted subcutaneously or even intravascularly. The primary reason for its attractiveness is the presence of substantive clinical results, which can in turn be credited to the long research history of this methodology. Indeed its principle of action is reasonably similar to the electrochemical sensing based blood glucose meters and therefore it inherits the inherent advantages of sensitivity, reproducibility and ease of fabrication. Nevertheless, implantation of the sensor *in vivo* presents a whole new set of challenges, as discussed below.

The development of subcutaneously implantable needle-type electrodes was first postulated by Schichiri and co-workers in 1985 [20]. Many design developments and clinical trials later, the

continuous glucose monitoring systems have reached a stable configuration consisting of a fine needle or flexible wire, with the active sensing element at or near the tip and implanted in the subcutaneous tissue. The functioning of the first-generation of electrochemical continuous glucose monitoring systems is based on the concentration sensitive reaction between oxygen and glucose catalyzed by glucose oxidase. The typical construction of these implantable sensors includes an electrode; an inner permselective membrane which rejects the interference molecules; an enzyme (glucose oxidase) layer; and an outer membrane which is highly permeable to oxygen but restricts the diffusion of glucose. The outer layer is also responsible for improving the biocompatibility of the device.

The second generation of these devices incorporates an exogenous mediator to react with glucose oxidase in place of oxygen [21, 22]. This modification allows the sensor to work independent of the (low) oxygen concentrations, which otherwise impede accurate glucose measurements. Clearly, a distinct parallel can be drawn between the advances made in blood glucose meter (test strip) technology and the electrochemical-based continuous glucose monitoring systems. However, a major challenge of employing the second-generation device *in vivo* is that the mediator must remain attached to the sensor. Additionally, the mediator must also retain its capability of enzyme reaction after it is coupled to the working electrode [23]. This problem was largely overcome by Dr. Adam Heller's invention of polymers that contained mediators and enzymes chemically attached to the polymer [24, 25]. The so-called "redox polymers" can react with glucose and, importantly, do not require oxygen. It is worth noting that this problem does not arise in blood glucose meter test strips as for the test strips, the mediator need not be confined to the working electrode.

At the time of this writing, three FDA-approved continuous glucose monitoring systems are available in the market, namely the Guardian Real Time (Medtronic), the DexCom STS (DexCom) and the FreeStyle Navigator (Abbott). On average, these continuous glucose monitoring systems display a prediction error percentage around 15%. As detailed in [26], the prediction accuracy depends on several factors including current glucose concentration and rate of change of glucose values, with poor correlation reported during hypoglycemia and times of rapid change (due to the physiological lag between blood and interstitial fluid glucose).

Despite its promising results, however, the current series of continuous glucose monitoring devices face multiple technical challenges in the road ahead. First, all of the currently approved sensors have a lifetime between 3-7 days inside the body. Long term stability of these devices is questionable and would need to be improved - both in terms of the ability to stay rooted to a specific region of the blood tissue matrix and in terms of the prediction robustness. Second, the ultimate goal of these devices is to combine their information with insulin pumps to form an "artificial pancreas" (often referred to as "closing the loop"). To that end, the sensor may need to be substantially more accurate if a real time insulin bolus is dependent on its readings. For example, one control algorithm for closing the loop states that the insulin bolus dose should be adjusted by 20% for glucose changes >2mg/dL [27]. Further, an adjustment of 10% insulin bolus is recommended for a change in glucose levels of the order of 1–2mg/dL. The current accuracy levels would fall short of meeting these prescribed diagnostic standards. Solution of these challenges is expected to facilitate tight glycemic control, thereby reducing potential diabetic complications significantly.

In addition to the aforementioned electrochemical *in vivo* sensors, considerable efforts have also been devoted to developing a reverse iontophoresis-based glucose sensing system. The

Cygnus GlucoWatch was based on this idea of extracting small amounts of interstitial fluid glucose using an electrical direct current (DC) through the skin [28]. The extracted glucose molecules were made to react with glucose oxidase in a gel collection disk leading to amperometric detection of the glucose concentration. While this device was approved by the FDA in 2001 (making it the second continuous glucose monitoring device to get FDA approval after the MiniMed CGMS), inaccurate readings, false alarms, local irritation on the insertion site with lack of improvement in glycemic control, and hypoglycemic episodes meant that it was withdrawn from the market after a short period of time.

2.3.2 Optical techniques

2.3.2.1 Fluorescence-based glucose detection

Motivated by the promise of a highly sensitive *in vivo* glucose sensing technique, a vast array of different techniques for fluorescence-based transduction of glucose levels has been investigated. The primary advantages of such an approach are the exquisite sensitivity (potentially extending to single-molecule detection levels) and the low cost of instrumentation for the full sensor device. In addition, these sensors afford complete electrical safety and the absence of electromagnetic interference, in contrast to the electrochemical sensors.

The central idea behind this sensing approach is to implant a small fluorescence-based device beneath the skin in a similar fashion to the electrochemical continuous glucose monitoring systems and to interrogate it by using an inexpensive optical probe from outside the body. This field was initiated by Schultz and co-workers with their pioneering efforts on the development of a glucose-binding protein, Concanavalin A (Con A), in a competitive binding assay with a high molecular weight dextran and an indwelling fiber optic component [29, 30, 31]. Since then,

research has been conducted on application of the resonance energy transfer technique [32] and fluorescence lifetime approaches [33] to this important diagnostic problem. The latter provides an attractive and reliable alternative to pure intensity measurements because fluorescence lifetime is essentially independent of any changes in light source intensity and detector sensitivity.

Currently, the principal foci of this exciting research avenue are on the design and synthesis of advanced glucose binding protein assays, including con A [34] and engineered glucose binding proteins [35]; enzyme-based glucose sensors that combine selective enzymes (such as glucose oxidase and apo-glucose oxidase) and optical indicators [36, 37]; and synthetic receptors such as boronic acid [38, 39]. A recent book edited by Geddes and Lakowicz [40] presents a fairly comprehensive review of the numerous fluorescence-based glucose detection methods that have been employed with varying degrees of success.

Despite its many advantages and rigorous results, however, no fluorescence-based glucose sensor has been commercially developed. Similar to the electrochemical subcutaneous glucose monitoring systems, long term stability needs to be improved for application in human subjects. Issues such as sensitivity in the UV-visible region in biological tissue as well as the modulation of the signal based on turbidity of the surrounding environment needs further investigation before one can draw a definitive conclusion on the in vivo viability of these sensors.

2.3.2.2 Carbon nanotube-based glucose detection

Recently, single walled carbon nanotubes (SWNT) have been employed as reporters for the optical detection of glucose by Strano and co-workers [41, 42]. The one-dimensional nature of the SWNT lends exquisite sensitivity to surface adsorption events. This salient feature coupled

with the discovery of near-infrared (NIR) fluorescence of SWNT in 2002 [43] has been instrumental in the development of these unique biosensors.

There are several advantages to employing the SWNT as optical reporters in place of the conventional fluorescent probes. First, their ability to fluoresce in the NIR region coincides with the diagnostic window (further detailed in Chapter 3), in contrast to the conventional fluorophores which largely emit in the UV-visible region of the electromagnetic spectrum. Furthermore, SWNT do not undergo photobleaching, which enables long-term use. On the other hand, typical organic fluorophores are susceptible to photobleaching, especially at high laser fluence.

Specifically, two important advances in this area have been reported. The first-generation sensor was based on the use of glucose oxidase in conjunction with a reaction mediator to control the SWNT fluorescence [41]. More recently, an affinity sensor based on SWNT was developed by employing competitive protein binding constructs. In this design, the SWNT is first coated in dextran, a glucose analogue. The addition of concanavalin A initiates protein-controlled aggregation and the fluorescence decreases as the number of free SWNT in solution is decreased. However, introduction of glucose leads to dissolution of glucose resulting in fluorescence recovery [42]. The ultimate goal of the researchers is to build on these concepts in making an *in vivo* sensor, which uses an appropriate coated dialysis capillary into which the SWNT sensing solution is loaded. The dialysis capillary would be freely permeable to glucose but not to the nanotubes contained inside. The total sensing device could then be implanted in a subcutaneous manner similar to the electrochemical continuous glucose monitoring systems. The optical setup would use a laser diode for excitation and a CCD camera for detection.

Notwithstanding the excellent results obtained over the past few years, multiple concerns remain about the *in vivo* viability of a nanotube based sensor. The first and foremost concern is that of biotoxicity. There is conflicting evidence in the literature on the cytotoxic properties of nanotubes and significant studies have to be accomplished to ensure that they are safe enough to be introduced into the human body, even if it is inside a dialysis capillary. Second, the accumulation and clearance inside the human body needs to be investigated in case of accidental leakage of the SWNT into the tissue. Third, the long term stability of such sensors would also play a crucial role, i.e. in terms of prediction robustness. Finally, tissue autofluorescence, while depleted, is not fully eliminated at the NIR wavelengths. Given the broad fluorescence signals, one aspect of investigation would be the potential uncertainty introduced into the measurements due to the spectral overlap between the SWNT and the tissue intrinsic fluorescence.

2.4 Non-invasive glucose sensing techniques

Based on our prior discussion of the different modalities, it is evident that a truly non-invasive measurement technique cannot have a chemical (or an electrochemical) component as such methods require some form of contact with the body, and often necessitate implantation underneath the skin surface. This restricts the scope of non-invasive glucose measurement technology, in which no sensor component is in contact with a biological fluid, to optical and spectroscopic methods only. Even within this relatively limited scope, a host of different methods have been proposed, a few of which are outlined in the following paragraphs. These methods have been selected based on their potential for performing glucose-specific measurements and on the strength of their published results.

2.4.1 Absorption spectroscopy

Among the different optical methods proposed for non-invasive glucose sensing, vibrational spectroscopy has shown the greatest promise. In particular, absorption spectroscopy - in the mid infrared (MIR) and especially in the near infrared (NIR) wavelength regime - have been extensively investigated for *in vivo* glucose sensing.

MIR absorption spectroscopy is used to interrogate the fundamental molecular vibrational modes in the range of 400-4000 cm^{-1} (2.5-25 μm). MIR absorption spectra are sharp and exhibit good spectral contrast between different bioanalytes. However, the high absorptivity of water in this spectral range results in penetration depths in biological tissue on the order of 10-100 μm. This means that MIR absorption based *in vivo* glucose sensing is not realistically feasible unless a separate methodology is incorporated for harvesting body fluid (e.g. the interstitial fluid). On the other hand, MIR absorption spectroscopy can be a useful tool for studying mixture samples *in vitro*, especially in the attenuated total reflection (ATR) configuration.

Shifting to the NIR regime (12500-4000 cm^{-1} or, 0.8-2.5 μm) solves the issue of penetration depth because of the presence of the so-called diagnostic window, due to a lack of significant absorbers in biological tissue in this wavelength region. This is in contrast to shorter wavelengths where hemoglobin absorption is substantial and longer wavelengths where water absorption predominates. Together with the low cost of typical instrumentation and large signal strength, this makes NIR absorption spectroscopy an attractive technology for *in vivo* glucose sensing.

However, as NIR light largely probes IR-active overtone or combination vibrational transitions, the resultant spectral bands tend to be significantly broader (i.e. there is a net loss in specificity in relation to MIR absorption or Raman spectra). Further, the implementation of NIR

spectroscopic methods has fundamental limitations due to the competing absorption by water, which though smaller than in the MIR region is not absent. The absorption by water also varies significantly as a function of temperature [44]. The application of multivariate calibration models presents a possible remedy for the determination of quantitative analyte-specific information.

For example, Heise *et al.* [45] performed a series of non-invasive NIR glucose measurements on a diabetic male volunteer. The spectral interval between 9000 and 5450 cm^{-1} was used to generate calibrations using partial least squares (PLS) regression. Their conclusion was that while this test provides evidence of glucose measurement, substantive improvement in the methodology is still required in the normoglycemic and hypoglycemic concentration ranges. Similarly, Samann *et al.* [46] evaluated the long-term accuracy of the NIR method by developing individual calibration models for 10 human subjects. These models were applied to spectra measured 84 to 169 days later and the mean square error of prediction varied from 55 to 646 mg/dL. This clearly illustrates the need for stable instrumentation and robust calibration algorithms to accommodate changes in patient physiology over the course of measurements. More recently, Olesberg *et al.* [47] performed in vivo NIR absorbance measurements on rat skin in the wavelength range of 2 to 2.5 μm during a glucose clamping experiment. Principal components were determined during a three hour euglycemic period. They then examined the change in absorbance during a 2 hour hyperglycemic period that is orthogonal to those principal components. Similarities were observed between these variations and the net analyte signal of glucose, suggesting the presence of glucose spectral information. In the industrial community, NIR System (formerly CME Telemetrix) measures absorbance of light transmitted through a finger. A broad spectrum between 500 and 1100 nm is collected and analyzed. This wavelength range is complemented with spectra in the 1100 to 1300 nm and 1590 and 1700 nm bands [48].

A cautionary note needs to be added to this section on NIR absorption spectroscopy based glucose sensing. As shown by Arnold and co-workers [49], the validity of published reports claiming to have successfully measured *in vivo* blood glucose from noninvasive near-infrared spectra collected in a time-dependent manner is questionable. Indeed apparently function models can be obtained when fictitious but time-dependent profiles were assigned to datasets containing no analyte information. One may conjecture that the non-specific, broad nature of the NIR absorption bands renders the developed calibration models susceptible to misinterpreting broad baseline drifts (resulting from temperature changes or system drift) as actual analyte features.

2.4.2 Near-infrared (NIR) Raman spectroscopy

NIR Raman spectroscopy combines the diagnostic window in the NIR region with the well-known fingerprinting ability of Raman spectroscopy. Moreover, modern improvements in Raman instrumentation allow low noise spectral data collection. The innate property of Raman spectroscopy, in all its forms, allows it to distinguish between molecules with great structural similarity (e.g., glucose and fructose) [50] making this technology an excellent candidate for glucose sensing, given the myriad other constituents present in the blood tissue matrix.

Our laboratory has played a pioneering role in the application of NIR Raman spectroscopy for *in vitro* and *in vivo* blood analyte detection. The results of our laboratory's efforts, with special emphasis on blood glucose monitoring, is detailed in Chapter 3.4. In this section, we briefly outline the reports of other investigators in this field.

Lambert *et al.* have demonstrated good predictability of Confocal Raman spectroscopy over a physiological range of glucose concentrations using PLS analysis in an anterior eye chamber model [51]. Glucose at clinically relevant concentrations was accurately predicted in human eye

aqueous humor *in vitro* using a PLS model based on artificial aqueous humor [52]. Progress toward the determination of glucose levels in patient-derived aqueous humor using laboratory-derived "artificial aqueous humor" calibration solutions were also reported [53].

Qu *et al.* described the use of Raman spectroscopy for noninvasive glucose measurements in human serum samples after ultra filtration [54]. In another investigation using serum and serum ultrafiltrate samples from 247 blood donors, the concentrations of glucose, triglycerides, urea, total protein, cholesterol, high density lipoprotein, low density lipoprotein and uric acid were determined with accuracy within the clinically interesting range. Concentrations were also predicted and relative errors of prediction around 12% were obtained. Differentiation between high- and low-density lipoprotein cholesterol, as well as the quantification of uric acid, was for the first time successfully accomplished for serum-based Raman spectroscopy [55]. Pilotto *et al.* [56] described the use of Raman spectroscopy to measure lactic acid, with the goal of measuring lactic acid in athletes undergoing exercise. Measurements were made in aqueous solutions, rat blood and *in vivo* in a rat.

In the context of *in vivo* glucose sensing in human subjects, Chaiken and co-workers [57] described using Raman spectroscopy to measure glucose from spectra measured on either the fingertip or the fingernail. Tissue modulation was used to reduce variations in tissue among individuals. The results show that even with this modulation, differences among individuals add significant error to measurements. The fluorescence spectra emitted from the same tissue region were also collected, with fluorescence intensity indicative of blood volume sampled [58]. This technology is being developed by Lightouch Medical. A later study in subjects with diabetes resulted in a correlation coefficient of 0.80 and an error of 22 mg/dL [59].

It is important to note that all of these reports deal with the application of *spontaneous* Raman spectroscopy. Recently, a series of publications from the van Duyne laboratory [60-64] have discussed the possibility of a surface enhanced Raman sensor for *in vivo* glucose sensing. The surface enhancement factor (typically in the range of 10^6 - 10^8) significantly reduces the problem associated with the low intrinsic sensitivity of Raman spectroscopy. Their current SERS sensor design relies on the glucose partitioning capability of a decanethiol/mercaptohexanol self-assembled monolayer and the surface enhancement effect of a silver film-over-nanospheres architecture.

In the *in vivo* studies conducted to date, a dorsal skinfold chamber has been implanted in an animal model (rat) for optical access to the SERS sensor. The ultimate goal of these studies is to develop an implantable SERS glucose sensor, similar in concept and construction to the fluorescent probe or SWNT-based glucose detection schemes. (As such, the SERS-based glucose sensing scheme is *not* non-invasive but has been included in this section for the sake of completeness.) Consequently, the SERS sensor inherits some of the same advantages and disadvantages of the other implanted sensor designs. In addition, one key concern remains the (quantitative) reproducibility of the glucose-specific signal, especially *in vivo*.

2.4.3 Optical activity and polarimetry

Polarimetry is a well-established technique for detection of analyte concentrations. This technique employs the ability of optically active molecules, such as glucose, to rotate the polarization plane of light. When a beam of plane-polarized light transverses a solution, its plane of polarization rotates by an angle, which is related to the concentration of the optically active solute, the sample path length and sample-specific rotation constant.

In 1982, March *et al.* first proposed the application of polarimetry for determination of blood glucose levels via the aqueous humor of the eye. They also postulated a potential *in vivo* system which would utilize a contact lens [65, 66]. Since this initial demonstration, the polarimetry technique has come a long way through the combined efforts of several investigators. Nevertheless, performance of polarimetry is largely restricted to the eye because scattering depolarizes light and the presence of other chiral molecules may confound measurements as well. In terms of *in vivo* sensing, Cameron *et al.* [67] presented glucose detection results from a laser-based optical polarimetry system using the anterior chamber of a rabbit eye. At this point in time, while the feasibility of predicting physiological glucose concentrations has been established *in vitro*, corneal rotation, corneal birefringence, and eye motion artifacts remain significant concerns in polarimetric ocular measurements.

2.5 References

(1) Khalil, O. S. *Clinical Chemistry* **1999**, 45, 165-177.

(2) Khalil, O. S. *Diabetics Technology & Therapeutics* **2004**, 660-697.

(3) Tuchin, V.V.; *Handbook of Optical Sensing of Glucose in Biological Fluids and Tissues* **2008**, CRC Press.

(4) Cunningham, D. D.; Stenken, J. A. *In vivo glucose sensing* **2009**, Wiley.

(5) T. M. Devlin (ed.), Textbook of biochemistry with clinical correlations, 6th edition, Wiley-Liss, N.Y., **2006**.

(6) Henry, J. B. Clinical diagnosis and management by laboratory methods, 19th ed.; Saunders: Philadelphia, **1996**.

(7) James, D.E.; Brown, R.; Navarro, J.; Pilch, P.F. Nature **1988**, 333, 183–185.

(8) Rother, K. I. The New England Journal of Medicine **2007**, 356 (15), 1499–501.

(9) De Fronzo, R. A. Diabetes **1988**, 37, 667-687.

(10) Kahn, C. R. Diabetes **1994**, 43, 1066-1084.

(11) Porte, D. Jr. Diabetes **1991**, 40, 166-180.

(12) The Diabetes Control and Complications Trial Research Group. New England Journal of Medicine **1993**, 329(14), 977–986.

(13) *Bode*, B.W.; *Schwartz,* S.; *Stubbs,* H. A.; *Block,* J. E. Diabetes Care **2005**, 28, 2361-2366.

(14) Wickramasinghe, Y.; Yang, Y.; Spencer, S. A. Journal of Fluorescence **2004**, 14(5), 513-520.

(15) Franciosi, M.; Pellegrini, F.; De Berardis, G.; Belfiglio, M.; Cavaliere, D.; Di Nardo, B.; Greenfield, S.; Kaplan, S. H.; Sacco, M.; Tognoni, G.; Valentini, M.; Nicolucci, A. Diabetes Care **2001**, 24, 1870-1877.

(16) Executive Summary: Standards of Medical Care in Diabetes—2010: Current criteria for the diagnosis of diabetes. Diabetes Care **2010**, 33, S4-S10.

(17) Kilpatrick, E. S.; Bloomgarden, Z. T.; Zimmet, P. Z. BMJ **2009**, 339, b4432.

(18) Turner, A. P. F.; Newman, J. D.; Tigwell, L. J.; Warner, P. J. Biosensors: A global view. The Ninth World Congress on Biosensors **2006**, Toronto, Canada, O31.

(19) Stenken, J. A.; Cunningham, D. D. Analytical Chemistry of in vivo Glucose Measurements, Wiley, **2008**.

(20) Schichiri, M. et. al. Future **1985**, 197-209.

(21) Habermuller, K.; Mosbach, M.; Schuhmann, W. Fresenius' Journal of Analytical Chemistry **2000**, 366, 560-568.

(22) Chaubey, A.; Malhotra, B. D. Biosensors and Bioelectronics **2002**, 17, 441-456.

(23) Henning, T. *Commercially available continuous glucose monitoring systems In in vivo Glucose Sensing. J. A. Stenken and D. D. Cunningham (eds).* Wiley, **2008.**

(24) Heller, A. *The Journal of Physical Chemistry* **1992**, 96 (9), 3579-3587.

(25) Csoeregi, E.; Quinn, C. P.; Schmidtke, D. W.; Lindquist, S. E.; Pishko, M. V.; Ye, L.; Katakis, I.; Hubbell, J. A.; Heller, A. *Analytical Chemistry* **1994**, 66 (19), 3131-3138.

(26) Cengiz, E.; Tamborlane, W. V. *Diabetes Technology and Therapeutics* **2009**, 11(S1), S-11–S-16.

(27) Wolpert, H. A. *Diabetes Care* **2008**, 31, S146–S149.

(28) Dunn, T. C.; Eastman, R. C.; Tamada, J. A. *Diabetes Care* **2004**, 27(9), 2161-2165.

(29) Mansouri, S.; Schultz, J. S. *Nature Biotechnology* **1984**, 2, 885-890.

(30) Schultz, J. S.; Mansouri, S.; Goldstein, I. J. *Diabetes Care* **1982**, 5, 245-253.

(31) Schultz, J. S.; Sims, G. *Biotechnology Bioengineering Symposium* **1979**, 9, 65-71.

(32) Meadows, D. L.; Schultz, J. S. *Analytica Chimica Acta* **1993**, 280, 21-30.

(33) Lakowicz, J.; Maliwal, B. *Analytica Chimica Acta* **1993**, 271, 155-164.

(34) Marvin, J. S.; Hellinga, H. W. *Journal American Chemical Society* **1998**, 120, 7.

(35) Tolosa, L.; Gryczynski, I.; Eichhorn, L. R.; Dattelbaum, J. D.; Castellano, F. N.; Rao, G.;Lakowicz, J. R. *Analytical Biochemistry* **1999**, 267 (1), 114-120.

(36) Trettnak, W.; Leiner, M. J. P.; Wolfbeis, O. S. *Biosensors* **1989**, 4(1), 15-26.

(37) Wolfbeis, O. S. *Sensors and Actuators B: Chemical* **1991**, 5, 1-6.

(38) Cordes, D. B.; Miller, A.; Gamsey, S.; Sharrett, Z.; Thoniyot, P.; Wessling, R.; Singaram, B. *Organic and Biomolecular Chemistry* **2005**, 3, 1708-1713.

(39) Cordes, D. B.; Suri, J. T.; Cappuccio, F. E.; Camara, J. N.; Gamsey, S.; Sharrett, Z.; Thoniyot, P.; Wessling, R. A.; Singaram, B. Two-Component Optical Sugar Sensing Using Boronic Acid-Substituted Viologens with Anionic Fluorescent Dyes in *Topics in Fluorescence Spectroscopy,* Geddes, C. D., Lakowicz, J. R., Eds., Springer, New York, **2006.**

(40) Geddes, C. D., Lakowicz, J. R. *Topics in Fluorescence Spectroscopy*, Springer, New York, **2006.**

(41) Barone, P.W.; Baik, S.; Heller, D. A.; Strano, M. S. *Nature Materials* **2005**, 4(1), 86-92.

(42) Barone, P. W.; Strano, M. S. *Angewandte Chemie International Edition* **2006**, 45(48), 8138-41.

(43) O'Connel,l M. J.; Bachilo, S. M.; Huffman, C. B.; Moore, V. C.; Strano, M. S.; Haroz, E. H.; Rialon, K. L.; Boul, P. J.; Noon, W. H.; Kittrell, C.; Ma, J.; Hauge, R. H.; Weisman, R. B.; Smalley, R. E. *Science* **2002**, 297(5581), 593-6.

(44) Langford, V. S.; McKinley, A. J.; Quickenden, T. I. *Journal of Physical Chemistry A* **2001**, 105(39), 8916-8921.

(45) Heise, H. M.; Bittner, A.; Marbach, R. *Clinical Chemistry and Laboratory Medicine* **2000**, 38(2), 137-145.

(46) Samann, A.; Fischbacher, C.; Jagemann, K. U.; Danzer, K.; Schuler, J.; Papenkordt, L.; Muller, U. A. *Experimental and Clinical Endocrinology & Diabetes* **2000**, 108(6), 406-413.

(47) Olesberg J. T.; Liu, L.; Zee, V. V.; Arnold, M. A. *Analytical Chemistry* **2006**, 78(1), 215-223.

(48) Pawluczyk, R.; Scecina, T.; Cadell, T. E. Method for determination of analytes using near infrared, adjacent visible spectrum and an array of longer near infrared wavelengths, *U.S. Patent*, No. 6741875, **2004**.

(49) Arnold, M. A.; Burmeister, J. J.; Small, G. W. *Analytical Chemistry* **1998**, 70(9), 1773-1781.

(50) Soderholm, S.; Roos, Y. H.; Meinander, N.; Hotokka, M. *Journal of Raman Spectroscopy* **1999**, 30(11), 1009-1018.

(51) Lambert, J. L.; Morookian, J. M.; Sirk, S. J.; Borchert, M. S. *Journal of Raman Spectroscopy* **2002**, 33 (7), 524-529.

(52) Lambert, J. L.; Pelletier, C. C.; Borchert, M. *Journal of Biomedical Optics* **2005**, 10(3), 1-8

(53) Pelletier, C. C.; Lambert, J. L.; Borchert, M. *Applied Spectroscopy* **2005**, 59(8), 1024-1031.

(54) Qu, J. N. Y; Wilson, B. C.; Suria, D. *Applied Optics* **1999**, 38(25), 5491-5498.

(55) Rohleder, D.; Kiefer, W.; Petrich, W. *Analyst* **2004**, 129(10), 906-911.

(56) Pilotto, S; Pacheco, M. T. T.; Silveira, L.; Villaverde, A. B.; Zangaro, R. A. *Lasers in Medical Science* **2001**, 16(1), 2-9.

(57) Chaiken, J.; Finney, W. F.; Yang, X.; Knudson, P. E.; Peterson, K. P.; Peterson, C. M.; Weinstock, R. S.; Hagrman, D. Progress in the noninvasive in-vivo tissue-modulated Raman spectroscopy of human blood, *Proceedings of SPIE*, 4254:216-227, **2001**

(58) Chaiken, J.; Peterson, C. Tissue modulation process for quantitative noninvasive in vivo spectroscopic analysis of tissues, *U.S. Patent*, No. 6289230, **2001**.

(59) Chaiken, J.; Finney, W.; Knudson, P. E.; Weinstock, R. S.; Khan, M.; Bussjager, R. J.; Hagrman, D.; Hagrman, P.; Zhao, Y. W.; Peterson, C. M.; Peterson, K. *Journal of Biomedical Optics* **2005**, 10(3), 031111.

(60) Shafer-Peltier, K. E.; Haynes, C. L.; Glucksberg, M. R.; Van Duyne, R. P. *Journal American Chemical Society* **2003**, 125, 588–593.

(61) Shah, N. C.; Lyandres, O.; Yonzon, C. R.; Walsh, J. T.; Glucksberg, M. R.; Van Duyne, R. P. *Analytical Chemistry* **2005**, 77, 6134–6139.

(62) Stuart, D. A.; Yonzon, C. R.; Zhang, X.; Lyandres, O.; Shah, N. C.; Glucksberg, M. R.; Walsh, J. T.; Van Duyne, R. P. *Analytical Chemistry* **2005**, 77, 4013–4019.

(63) Stuart, D. A.; Yuen, J. M.; Lyandres, N. S. O.; Yonzon, C. R.; Glucksberg, M. R.; Walsh, J. T.; Van Duyne, R. P. *Analytical Chemistry* **2006**, 78, 7211–7215.

(64) Yonzon, C. R.; Haynes, C. L.; Zhang, X. Y.; Walsh, J. T.; Van Duyne, R. P. *Analytical Chemistry* **2004**, 76, 78–85.

(65) Rabinovitch, B.; March, W. F.; Adams, R. L. *Diabetes Care* **1982**, 5, 254–258.

(66) March, W. F.; Rabinovitch, B.; Adams, R. L. *Diabetes Care* **1982**, 5, 259–265.

(67) Cameron, B. D.; Gorde, H. W.; Satheesan, B.; Cote, G. L. *Diabetes Technology and Therapeutics* **1999**, 1, 135-143.

CHAPTER 3

AN INTRODUCTION TO RAMAN SPECTROSCOPY

Raman spectroscopy is a fundamental form of molecular spectroscopy that is widely used to investigate the structures and properties of molecules from their vibrational transitions [1]. The spectroscopic technique itself is based on the phenomenon of Raman scattering, an inelastic scattering process stemming from the interaction of light with matter, which was postulated by Smekal in 1923 [2] and first observed experimentally by Raman and Krishnan in 1928 [3]. During the interaction process, energy is transferred either to (Stokes) or from (Anti-Stokes) the vibrational or rotational modes of the molecule. Consequently, information on the molecular structure is encoded in the corresponding frequency shifts of the scattered photons (red-shifted for Stokes and blue-shifted for Anti-Stokes) that appear in the form of characteristic spectral patterns ('fingerprinting').

Despite the discovery of the Raman effect in the early twentieth century, most of the initial work was confined to physical and structural investigations, with relatively few reports on chemical and biological analysis. However, remarkable improvements in Raman instrumentation (e.g. efficient spectrometers, cooled multichannel detectors and holographic filters), particularly over the last two decades, has led to a period of intense activity with a paradigm shift in its application areas to chemical process monitoring and biomedical research. Specifically, investigators have exploited its unique fingerprinting capabilities to study the structure and compositional changes in biological tissue for diagnosis of disease *in vitro* and *in vivo* [4-7].

In this chapter, we first outline the theory of Raman scattering and its relevant interpretations. Subsequently, we discuss about the instrumentation needs for acquiring high quality Raman spectra, with special emphasis on biological samples. The application of multivariate calibration modeling to analysis of spectroscopic datasets is then detailed. This sets the stage for discussions on the application of Raman spectroscopy for blood glucose monitoring, especially the past efforts of this laboratory. Finally, we conclude by inferring the technical challenges from the results of our previous studies. The solution of these challenges forms the core of this thesis (Chapter 4-8).

At this juncture, it is worth emphasizing that this thesis is based on the applications of *spontaneous* Raman scattering and all discussions on Raman spectroscopy pertain to this effect only, unless otherwise qualified.

3.1 Theory of Raman scattering

During interaction of light with matter, the photons of the incident beam can either be absorbed, scattered or transmitted without interaction. In turbid biological media, the first two processes are predominant. For the absorption process, the energy of the incident photon must match the energy gap between the relevant vibronic states of the molecule (i.e. ground state and excited state) resulting in the promotion of the molecule to the higher energy excited state. This principle is employed in absorption spectroscopy, such as NIR absorption spectroscopy where infrared energy covering a range of frequencies is directed onto the sample and the loss of specific frequencies in the acquired signal is interpreted as the difference between the ground state(s) and the corresponding excited state(s) of the vibrational mode(s) of the molecule(s). If the molecule is excited to a higher electronic state, other subsequent processes such as fluorescence emission

and non-radiative transitions (including dissipation of energy from the molecule to the surroundings) may be observed.

In contrast, the scattering process does not need the incident photon to have a specific energy matching the difference in energy levels of the molecule. In classical terms, the scattering phenomena can be described as a perturbation of the electronic cloud of the molecule due to the interaction with the incident electric field and the resultant modulation of the electric field by the vibrating molecule. As described mathematically later in this section, the modulation results in two types of scattered radiation, namely at the original frequency of the incident beam (Rayleigh scattering) as well as at frequencies different from that of the incident beam (Raman scattering). It is evident that for Rayleigh scattering, there is no net change in the vibrational energy of the interrogated sample, while for Raman scattering the incident radiation can either transfer energy to the molecule (resulting in lower frequency of the scattered radiation, Stokes scattering) or gain energy from it (resulting in higher frequency of the scattered radiation, anti-Stokes scattering).

It is important to note that the scattering process occurs via an intermediate "virtual state" (Fig. 3.1), which is not necessarily a true quantum state of the molecule but can be considered a very short lived distortion (typically 10^{-14} second or less) of the electron cloud caused by the oscillating electric field of the light. In other words, Raman scattering can be observed for any excitation frequency through the formation of an appropriate virtual state. Consequently, the acquired Raman spectra contain specific peaks that are characteristic of the molecular structure, but not a function of the excitation frequency (although the intensity of the radiation has a strong dependence on the excitation frequency).

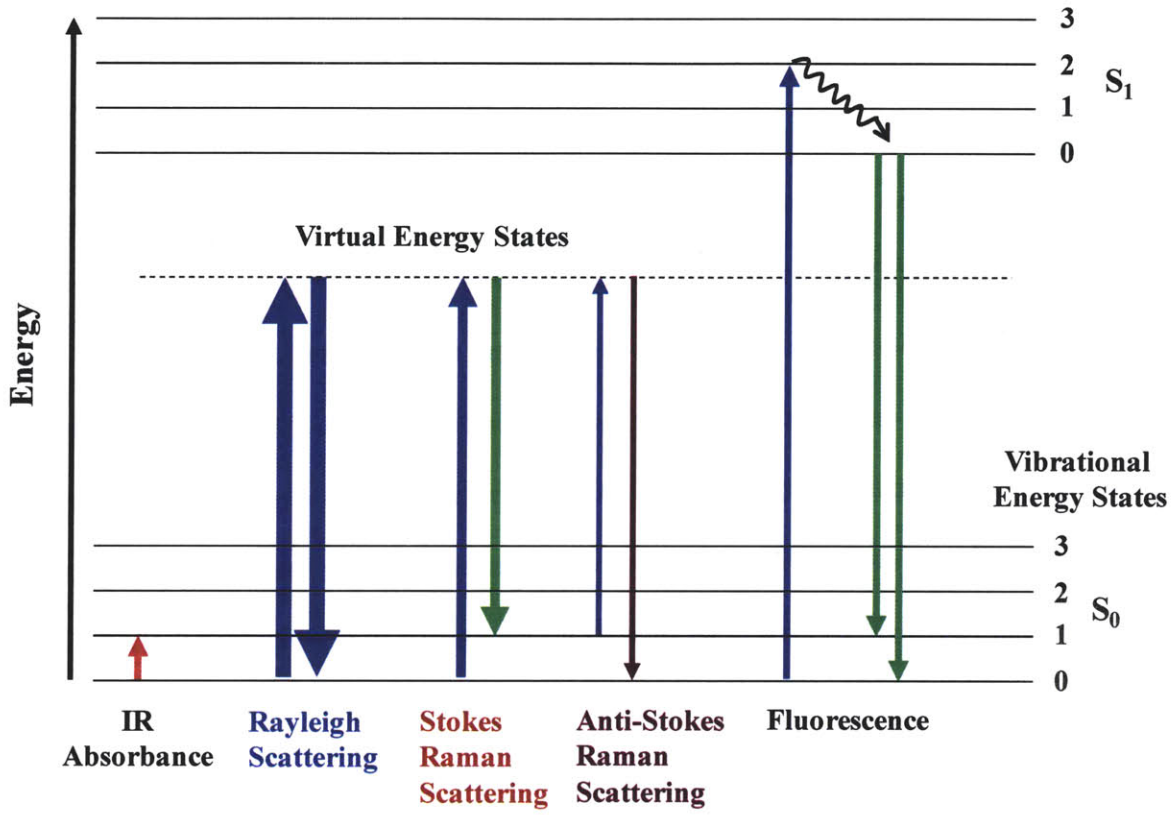


Fig. 3.1 Jablonski diagram depicting the transitions between the different vibronic states as a result of several light-matter interactions. Here, S_0 and S_1 represent the ground electronic and excited electronic states, respectively.

The Jablonski diagram, shown in Fig. 3.1, is a useful tool for illustrating the vibrational and electronic states and the aforementioned transitions between them (IR absorption, fluorescence, Rayleigh scattering, Stokes scattering and anti-Stokes scattering). Here, radiative transitions are depicted by straight arrows whereas nonradiative transitions (such as the vibrational relaxation step in the fluorescence emission process) are given by squiggly arrows.

In general, most of the photons are elastically scattered and only one in every 10^6-10^8 photons undergo Raman scattering. Also, Stokes scattering is typically much stronger than anti-Stokes under room temperature conditions, since the number of molecules in an excited

vibrational state is significantly smaller compared to those in the low-energy ground state (as per the Boltzmann equation). Therefore, most Raman spectra for biological samples is acquired only on the low-energy side, although this has the disadvantage of being intrinsically coupled to the fluorescence interference (which always occurs at a lower energy than the excitation frequency as shown in Fig. 3.1).

Using the classical picture previously alluded to, one can formulate a simple mathematical model of the scattering process by studying the interaction between an incident electromagnetic wave and the sample's molecular/atomic structure. In the following, we develop such a model to provide greater insight and context to our previous discussion of the scattering phenomena. (For a more detailed theoretical background, the reader is referred to several excellent references in the literature [8, 9].)

As the incident radiation (i.e. an electromagnetic wave in the classical sense) interacts with the molecules of the sample, the electron orbits within the constituent molecules are perturbed periodically with the same frequency (v_0) as the excitation frequency. The consequent perturbation or distortion of the electron cloud in relation to the nucleus results in a periodic separation of charge within the molecules, which is called an induced dipole moment.

This induced dipole moment (P) is proportional to the polarizability (α), i.e. the relative tendency of a charge distribution to be distorted from its intrinsic shape by an external electric field, and the incident electric field intensity (E_i):

$$P = \alpha E_i . \tag{3.1}$$

The polarizability is a material property that depends on the molecular structure and nature of the bonds holding the atoms within the molecule. The electric field component of the incident radiation E_i, at a certain time t, can in turn be expressed as:

$$E_i = E_0 \cos(2\pi\nu_0 t), \tag{3.2}$$

where E_0 is the amplitude and ν_0 is the excitation frequency. Combining Eq. (3.1) and Eq. (3.2), we get:

$$P = \alpha E_0 \cos(2\pi\nu_0 t) \tag{3.3}$$

Additionally, assuming sinusoidal pattern of vibrations for the molecules of the sample, we can write the small nuclear displacements (dx) about the equilibrium position as:

$$dx = x_0 \cos(2\pi\nu_{vib} t), \tag{3.4}$$

where x_0 is the vibration amplitude and ν_{vib} represents the frequency of intrinsic vibration of the molecular mode.

Further, given the dependence of the polarizability on the instantaneous position of the atoms (as this directly affects the ability of the electric field to perturb the electron cloud), one may write the instantaneous polarizability in the form of a Taylor expansion for small nuclear displacements as:

$$\begin{aligned} \alpha &= \alpha_0 + \frac{\partial \alpha}{\partial x} dx \\ &= \alpha_0 + \frac{\partial \alpha}{\partial x} x_0 \cos(2\pi\nu_{vib} t) \end{aligned}, \tag{3.5}$$

where α_0 is the polarizability of the molecular mode at equilibrium position. (Higher order terms are neglected because such nuclear displacements are usually a very small fraction of the characteristic length scale, i.e. the bond length.)

Plugging Eq. (3.5) in Eq. (3.3) and adjusting the two cosine terms using a trigonometric identity, we obtain:

$$P = \alpha_0 E_0 \cos(2\pi\nu_0 t) + \frac{\partial\alpha}{\partial x}\frac{x_0 E_0}{2}\left[\cos(2\pi(\nu_0 - \nu_{vib})t) + \cos(2\pi(\nu_0 + \nu_{vib})t)\right] \qquad (3.6)$$

Equation (3.6) shows that the induced dipole moment is created at three different frequencies, namely ν_0, ν_0-ν_{vib} and ν_0+ν_{vib}. Since the induced dipole moment is in turn manifest as a source of the scattered electromagnetic wave, it is evident that the scattered radiation also results at these three frequencies, i.e. Rayleigh scattering, Stokes Raman scattering and anti-Stokes Raman scattering, respectively. .

In addition, Eq. (3.6) also implies that Raman scattering is possible if and only if the derivative of polarizability with respect to the displacement, $\partial\alpha/\partial x$, is non-zero. In other words, the condition that the polarizability must change with vibrational displacement about the equilibrium position provides a gross selection rule for Raman-active vibrations. It is worth mentioning that polarizability usually varies with spatial direction (anisotropic property of a molecule) and is independent of the permanent dipole moment, if present, in a molecule. The latter is of significance as it provides the analogous selection rule for infrared absorption spectroscopy where absorption of a specific mode correlates to the corresponding change in the permanent dipole moment of the molecule. Evidently, not all vibrations of the molecule are both Raman and IR active typically resulting in different characteristic spectra, although some common peaks may be observed (except for centrosymmetric molecules where there is mutual

exclusivity between the bands obtained by the two techniques). The two methods are, therefore, often viewed as providing complementary information.

Finally, it can be shown that the Raman scattering intensity is proportional to the incident intensity, square of the polarizability and, most importantly, the fourth power of $v_0 \pm v_{vib}$ [10, 11]. This implies Raman scattering is more efficient at higher laser frequencies (lower wavelengths), which of course would need to be balanced with several considerations for the specific sample type (e.g. fluorescence and presence of significant absorbers). Further, since the Raman intensity also depends on the square of the polarizability, molecules with distributed electron clouds, such as those containing double bonds, typically tend to be strong Raman scatterers. This can be primarily attributed to the fact that such electron clouds, e.g. the π-electron cloud of the double bond, are easily distorted on application of an external electric field.

While the classical formulation presented here provides some valuable insights into the essential concepts of Raman scattering, it presents an incomplete picture of the true process. For example, the (different) intensities of Stokes and anti-Stokes scattering (or, resonance Raman scattering) cannot be directly quantified using the classical theory. A quantum mechanical approach, where Raman scattering is detailed using the Kramers Heisenberg Dirac (KHD) equation, is necessary for such cases. The elucidation of the KHD equation is beyond the scope of this current thesis but the interested reader is referred to [12].

3.2 Instrumentation

As was previously alluded to, the singular obstacle to the widespread usage of Raman spectroscopy for biological applications was the dearth of suitable instrumentation. Over the years, however, this problem has been largely overcome with a mixture of scientific ingenuity

and critical engineering advances. In the following, we outline the considerations for performing Raman spectroscopy on biological materials and the experimental systems that are used to acquire the spectra. Further, we interpret the acquired Raman spectra for specific analytes in the context of the Raman scattering theory presented in Chapter - 3.1.

The primary consideration for performing biological Raman spectroscopy lies in the selection of a suitable excitation wavelength. Traditionally, for chemical analysis UV or visible wavelengths are employed, mainly because of the fourth power dependence of the Raman scattering intensity on the excitation frequency. Moreover, the efficiency of detectors available in the UV-visible region is significantly better than those available in the NIR region. Specifically, silicon-based detectors are not possible beyond 1100 nm (due to the intrinsic Si band gap of 1.1 eV) and their quantum efficiency falls rapidly for detection above 1000 nm.

Nevertheless, NIR excitation provides some crucial advantages for *in vitro* and *in vivo* characterization of biological tissue. First, the fluorescence interference is substantially reduced by shifting the excitation from UV-visible to NIR wavelengths. This is because the primary chromophores in tissue absorb and emit in the UV-visible region, which means only a broad 'tail' of the fluorescence signal remains in the NIR region. Second, thermal and photochemical damage are minimized by the use of low energy NIR light. NIR wavelengths are not mutagenic (especially as compared to UV wavelengths) and absorption is relatively low so tissue heating is also minimized. Third, working in the NIR region, especially 700-1000 nm, provides substantial penetration depth into biological tissue because of the lack of significant absorbers in this wavelength region. The relatively large penetration depth (~1mm) of this wavelength region, sometimes called the diagnostic window [13], provides the opportunity to observe Raman information of underlying components, such as blood analytes.

As a result, for most tissue optics studies, and especially for biological Raman spectroscopy, an excitation wavelength in the NIR region is chosen. The commonly used wavelengths are 785 nm and 830 nm, the merits of which could be balanced for a particular application based on the criteria listed above. The use of such a wavelength also entails special modifications to the other optical elements particularly the CCD used for spectra acquisition (discussed below).

Generally, an experimental system for Raman spectroscopy has the following major components: (a) an excitation source, usually a continuous-wave (CW) laser; (b) sample excitation and collection system; (c) a dispersive element coupled with the detector. We briefly review each of these components in the following paragraphs. It is noted that the scope of our discussion will be restricted to dispersive Raman spectrometers, given their much greater prevalence in biological applications. The other primary type of spectrometers, namely Fourier transform (FT)-Raman spectrometers [14], provide excellent frequency precision and better fluorescence avoidance (at higher NIR excitation wavelengths) but suffer from significant detector noise in comparison to the CCD detectors employed for the dispersive systems [10].

As discussed before, in biological Raman spectroscopy, NIR excitation at 785 or 830 nm is most common. Depending on the type of application (e.g. confocal Raman microscope vis-à-vis clinical Raman instrument), the flexibility of the excitation source has to be balanced with its spatial footprint, or compactness. In a confocal Raman microscope, such as the one built in our laboratory for monitoring of intra-cellular processes, a titanium sapphire laser pumped by a solid state laser may be used as it offers a dynamic tunable excitation range that can be optimized for different applications. Nevertheless, its size and cost prohibits its application in clinical Raman instruments where a smaller and cheaper external cavity solid state diode laser is preferred. In both our benchtop and clinical blood glucose monitoring systems, an 830-nm external cavity

diode laser (Process Instruments) is employed as the Raman excitation source. It is worth mentioning that the cheaper lasers (due to imprecise temperature control) are prone to exhibiting amplified spontaneous emissions that require an additional laser line filter for suppression.

The next critical component of the Raman system is the design and assembly of the excitation and collection optics. The system maybe designed for free space optics, fiber optics or a combination of the two. Usually, fiber optics is preferred for clinical instrumentation due to the inherent flexibility and compactness. However, for bench-top devices free space optical elements may be able to provide greater collection efficiency. In particular, non-imaging optical elements such as gold-coated off-axis paraboloidal mirrors [15] and compound parabolic concentrators (CPC) [16] have been incorporated to enhance the effective numerical aperture of the collection optics. Traditionally, most biological Raman spectroscopy systems are operated in the 180° back-scattering configuration, primarily because the turbidity prevents sufficient photons from being collected in other modes. Nevertheless, recently other variations including 90° collection geometry, transmission geometry and spatially offset Raman have gained a certain measure of popularity [17, 18]. This aspect is further discussed in Chapter 9.

In our bench-top (reflection mode) Raman system, the excitation beam is guided through a hole in the paraboloidal mirror to the sample. The sample is aligned such that most of the specular reflection escapes through this same hole. The remaining back-scattered light is then collected and directed by the paraboloidal mirror towards a holographic notch filter for suppression of the Rayleigh scattered light. The filtered beam is then focused onto an optical fiber bundle which serves as a conduit to the spectrograph slit. Importantly, all optical elements such as focusing lens and steering mirrors in such a free-space optical system should be NIR anti-reflection coated to achieve optimal performance.

For our clinical system, we employ a specialized probe for sample excitation and collection. This probe, similar to the one designed by Motz *et al.* [19], features a central excitation fiber surrounded by a ring of collection fibers. The number and diameter of these fibers as well as the number of annular rings is determined by *a priori* optimization of the radial and angular collection efficiency. The height of the spectrograph slit provides an additional constraint on the parameters of the collection fibers. In our current system, we employ a single ring of 9 low-OH collection fibers (200 μm ID, 220 μm OD). The fiber probe is capped by a shell and tube miniaturized filter assembly, which functions to suppress the Raman bands produced in the central excitation fiber as well as the reflected elastically scattered light in the surrounding collection fibers. The schematics of the bench-top Raman system and the clinical unit are shown in Fig. 3.2.

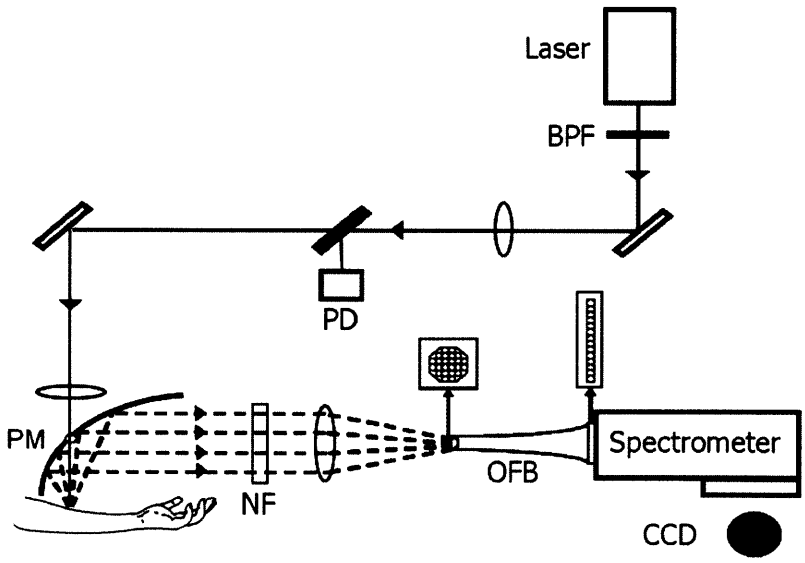


Fig. 3.2(A) Schematic of bench-top Raman system for transcutaneous blood glucose detection. BPF: bandpass filter; PD: photodiode; PM: paraboloidal mirror; NF: notch filter; OFB: optical fiber bundle.

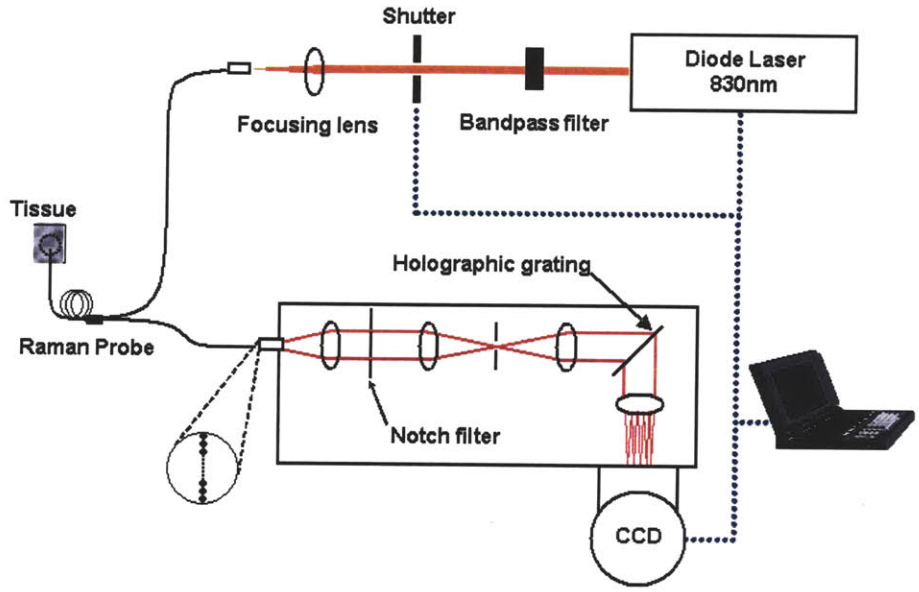


Fig. 3.2(B) Schematic of clinical Raman system for transcutaneous blood glucose detection.

In dispersive Raman spectrometers, the critical element is the grating which can simultaneously diffract different wavelengths of the input beam onto a multichannel detector. In other words, the wavelength distribution is transformed to a linear spatial distribution by the dispersive element. Holographic gratings are commonly employed for this purpose due their intrinsic advantage of very low stray light from surface scattering, although their efficiencies are usually lower than those of mechanically ruled gratings. In our Raman systems, we employ holographic grating-based spectrographs with f/# ranging from f/1.4 to f/2.0, depending on the specific application. Finally, the dispersed light is collected using a cooled CCD (charge coupled device). Deep depletion, back-illuminated CCD is the device of choice as they provide significantly higher quantum efficiency than their front-illuminated counterparts and reduce the etaloning effect for NIR wavelengths. To reduce the dark current to acceptable levels, especially given the limited sensitivity of Raman signals, state-of-the-art liquid nitrogen or Peltier cooling is used. The liquid-nitrogen cooled detectors provides better dark current characteristics and are

ideally suited for laboratory systems, while the thermoelectrically cooled CCD are incorporated in our clinical Raman systems due to their compactness and convenience.

3.3 Data interpretation and modeling

3.3.1 Data Interpretation

A Raman spectrum is usually plotted with the scattered intensity (counts measured by the CCD) on the y-axis and the energy on the x-axis, as shown for a calibration standard (acetaminophen powder) in Fig. 3.3. The energy on the x-axis is expressed in terms of the Raman shift ($\Delta\tilde{\nu}$) in wavenumbers, i.e. (ν_0-ν_{vib})/c in units of cm^{-1}. (For purposes of brevity, several authors simply state the Raman shift as $\tilde{\nu}$.) Historically, wavenumber presents a convenient spectroscopic quantity for spectral analysis as it is directly proportional to the frequency (or difference in energy levels) and also because spectroscopic instruments are typically calibrated in terms of wavelength (1/wavenumber), independent of the speed of light or Planck's constant.

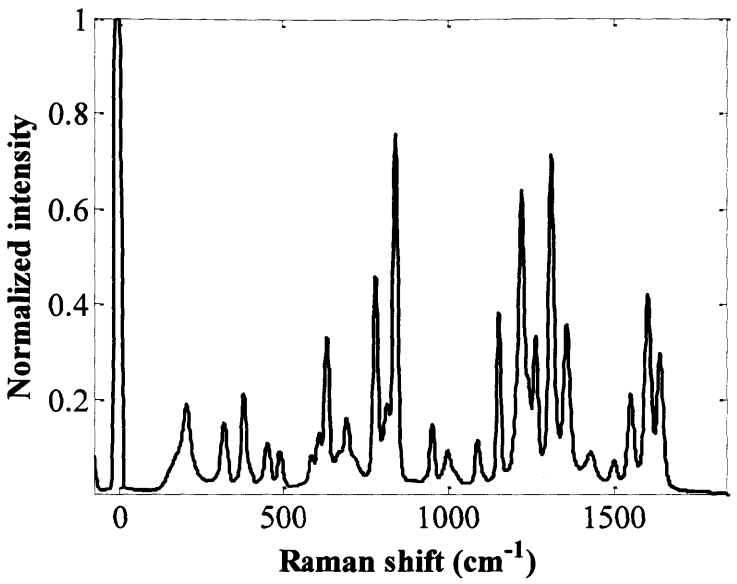


Fig. 3.3 A Raman spectrum of acetaminophen powder, which is often used as a calibration standard [20].

It is notable that the Rayleigh line at 0 cm^{-1} shows the highest intensity even after the notch filter action. Since the optical density of a typical notch filter is better than 6 (i.e. it provides a factor of 10^6 reduction in the intensity of the specific line), the Raman intensity evidently is a very small fraction of the elastically scattered light. This observation is unsurprising, even though acetaminophen powder is a very strong Raman scatterer. As a consequence of this and the potential imprecision associated with the notch filter action, Raman spectral analysis is rarely performed on the 0-200 cm^{-1} region.

Figure 3.4 shows the Raman spectra acquired from the aqueous solutions of three clinically relevant blood analytes, namely glucose, urea and creatinine. The spectra were acquired using our laboratory system from aqueous solutions of these analytes held in a fused silica cuvette. The three spectra show the excellent specificity of this analytical technique. Indeed, the innate

property of Raman spectroscopy allows it to distinguish between molecules with great structural similarity (e.g. glucose and fructose) [21]. Furthermore, it is evident that the water contribution to these signals is negligible, which is in stark contrast to the observations in NIR absorption spectroscopy.

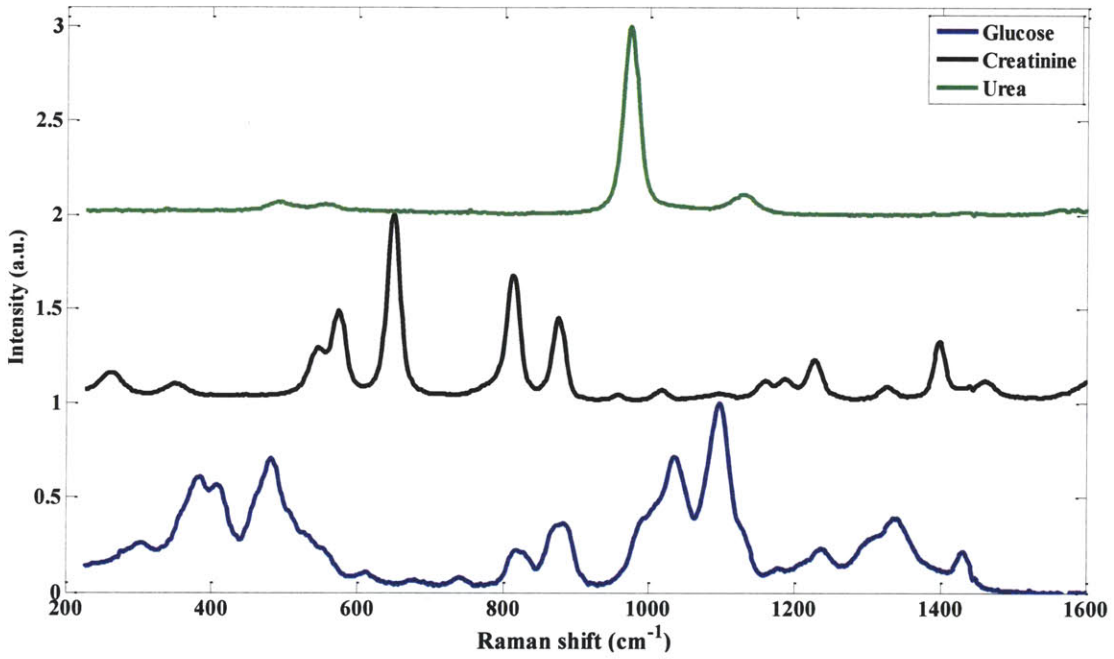


Fig. 3.4 Raman spectra of anomeric balanced D-glucose (aq.), creatinine (aq.), and urea (aq.) (spectra are offset for clarity).

One can, of course, investigate the molecular structure and properties based on the Raman spectra acquired from these analytes. This forms the core of the structural investigations, which makes Raman spectroscopy an important tool in physical chemistry. For example, the glucose Raman spectrum shows characteristic molecular vibrations including the δ(C-C-C), δ(C-C-O), δ(C-O) and τ(C-C) skeletal modes between 380-500 cm^{-1}; υ(C-O), υ(C-C), δ(C-C-H), δ(C-C-O) stretch between 820-950 cm^{-1}; υ(C-C), υ(C-OH), δ(C-O-H) stretch between 1000-1180 cm^{-1}; and δ(O-C-H), δ(C-C-H) , CH_2, δ(C-O-H) bending between 1200-1500 cm^{-1} [21, 22]. (The specific

structural references for the glucose molecule are provided in Fig. 2.1.) In this thesis, however, we will primarily focus on the analytical capabilities of the Raman spectra, i.e. the ability to provide quantitative predictions of concentration information of Raman-active analytes in complex tissue matrices. In this context, we conclude our discussion of the interpretation of the Raman spectra with a brief note on the widely used empirical relationship between the analyte concentration and the relevant Raman scattering intensity.

While the theoretical (both the classical and the quantum mechanical) treatment of the Raman scattering process (Chapter-3.2) provides a detailed understanding of the various subtleties involved, in most Raman experiments the only parameter of interest is the signal strength. Thus, a detailed correlation of the experimental data and the theoretical predictions is usually not performed. A more pragmatic approach, based on a simple linear relationship between the relevant analyte parameters and the acquired dataset, is adopted. To this end, Raman spectroscopists define an empirical parameter, the differential molecular Raman scattering cross-section (β) at a particular excitation wavelength, to indicate the probability of an incident photon being scattered as a Raman-scattered photon for a given Raman shift. One can then write the following equation in a manner analogous to Beer's law for absorption [10]:

$$P_{Raman} = P_0 \beta c dz, \tag{3.7}$$

where P_{Raman} represents the Raman scattering intensity in a steradian of collection solid angle (photons cm^{-2} sr^{-1} s^{-1}), P_0 is the incident laser intensity (photons cm^{-2} s^{-1}), c is the number density of scatterers (concentration in molecule cm^{-3}), and dz is the path length of the laser in the sample (cm). Clearly, β has units of cm^2 sr^{-1} $molecule^{-1}$. It is instructive to note that cdz denotes the total number of excited molecules for a beam having a cross-sectional area of 1 cm^2.

Additionally, Eq. (3.7) can be further re-written to define a Raman scattering cross-section (μ_{Raman}), a functionally similar parameter to the elastic scattering (μ_s) and absorption (μ_a) coefficients:

$$\mu_{Raman} = 4\pi\beta c \tag{3.8}$$

Previously, investigators have reported that the differential Raman scattering cross-section for aqueous glucose solution is 5.6×10^{-30} cm^2 sr^{-1} $molecule^{-1}$ [10]. This value was measured for the Raman peak at 1126 cm^{-1} using laser excitation frequency of 514 nm. Using this value, one can readily compute that for typically observed values of physiological glucose concentrations (80-400 mg/dL), the Raman scattering cross-section (μ_{Raman}) lies in the range of 10^{-10}-10^{-9} cm^{-1}. This number can then be compared to typical values of 40 cm^{-1} for μ_s and 0.1 cm^{-1} for μ_a in biological tissue. While these numbers would be slightly different for our NIR measurements, they provide important insights into the relative probability of the different vibronic processes. Clearly, a miniscule fraction of the light is Raman scattered and this low sensitivity presents substantive challenges for quantitative predictions, despite the excellent chemical specificity of Raman spectroscopy.

3.3.2 Modeling and calibration

Data analysis forms one of the central tasks in transcutaneous studies [23], because the Raman signature of glucose *in vivo* is very weak, as shown above. In addition, the presence of interfering chemical components (e.g. other Raman-active tissue constituents such as proteins and lipids) pose a significant challenge to extracting glucose-specific information. As a result, many different analytical techniques, more properly called chemometric techniques, have been proposed and used with varying degrees of success for Raman spectral analysis. (Chemometrics

can be formally defined as the chemical discipline that uses mathematical and statistical methods, (a) to design or select optimal measurement procedures and experiments, and (b) to provide maximum chemical information by analyzing chemical data.)

In this section, we outline the fundamental ideas of the relevant chemometric calibration procedures and their respective merits. The emphasis here is on introducing the reader to the basic principles of calibration, rather than dwelling on the intricate mathematical schemes, in order to remove the "black-box" mystique that often shrouds its application. The reader is referred to excellent articles in the literature for a more exhaustive treatment [24-26].

Loosely speaking, the gamut of spectral data analysis ranges from using numerical methods to improve data quality (e.g. effective signal-to-noise ratio), application of the regression or classification schemes to extract the information of interest from the spectral dataset and finally the validation and visualization of the information content [27]. The first falls under the general class of pre-processing method, while the latter two are considered to be post-processing.

3.3.2.1 Pre-processing methods

Pre-processing methods are often key to interpreting the scientific content of the data structure. Such methods range from simple spectral corrections to sophisticated algorithms that guarantee and enhance data quality as well as provide consistency for the subsequent processing steps. Specifically, for biomedical spectroscopy given the data flood often produced from a single experiment, the aspect of ensuring consistency in the datasets that are inputted for further calibration or classification analysis cannot be emphasized enough.

The Raman spectra acquired in our laboratory, whether using the bench-top laboratory system or the clinical Raman unit, are subjected to the same set of pre-processing steps to ensure

reproducibility. First, because of the unusually large area CCD employed in our system (typically 1" by 1"), the images obtained on the chip display a curved profile instead of a linear one. This can be readily visualized by looking at the Rayleigh scattered light entering the spectrograph from the linear array of fibers. As a result, curvature correction is employed before standard vertical binning [20], failing which one may obtain a net loss in the spectral resolution. After a binned (one-dimensional) spectrum is obtained, wavenumber calibration and spectral response correction is performed. Since the Raman shift in wavenumber does not have a linear relationship with the spectral dispersion (unlike wavelength), a standard method is to acquire the Raman spectrum of a well-characterized Raman scatterer (e.g. calibration standard such as acetaminophen powder) and employ a polynomial interpolation between the known bands and the CCD pixels to provide a full axis wavenumber calibration. This axis can then be used for all other acquired spectra. The sample Raman spectrum is then corrected for the intrinsic system response by normalizing with a spectrum acquired from the system using a industry standard wavelength-calibrated white light source.

Another important step is removal of spike effects, which are introduced by high energy cosmic rays or or γ- and α-rays emitted from materials around the detector chip. Several methods can be used for the suppression of such spikes, including median filtering (or some derivative of the same) by comparison of multiple (temporal or spatial) frames and band width-based threshold filtering (based on the significantly larger linewidths of the analyte spectrum as compared to the spike) [28].

In addition, depending on the particular application, one may employ baseline correction, spectral smoothing schemes and de-noising algorithms. The key is to critically investigate the usefulness of these techniques in terms of providing improvements in predictive ability in

relation to the possibility of incorporation of spurious artifacts in the acquired spectrum. For example, while baseline correction has been pursued extensively in Raman spectroscopy, in most applications implicit calibration techniques such as partial least squares (PLS) exhibit equally good, if not superior, performance without this additional pre-processing step (as discussed in Chapter 6) [29]. Similarly, smoothing schemes, such as the Savitzky-Golay smoothing algorithm [30], by themselves do not provide substantive benefits for our analysis. Nevertheless, in specific applications, e.g. analysis of first derivative Raman spectra, the smoothing schemes provide a tremendous boost.

3.3.2.2 Post-processing methods

Chemometrics encompasses a wide variety of analysis and experimental design methods, but there are two principal application areas which stand out in terms of their significance in biomedical spectroscopy. These two applications, namely calibration (or regression) and classification (or pattern recognition), arise in different categories of biomedical spectroscopy problems. Calibration is generally encountered when a spectroscopic measurement must be used to predict the value of an underlying parameter (e.g. analyte concentration) or property. On the other hand, classification schemes involve the use of the spectral dataset to determine patterns in the biomedical context (e.g. histo-pathological classes for cancer detection) and to use this information to provide prospective discrimination. Evidently, the blood glucose monitoring problem falls under the broad set of calibration applications and in our ensuing discussions we will focus on the calibration portion of chemometrics investigation. Nevertheless, it is worth mentioning that the formulation of classification and regression schemes themselves are often intertwined. For example, a partial least squares (PLS) calibration scheme can be readily adapted for partial least-squares-based discriminant analysis, PLS-DA (a classification algorithm).

Conversely, support vector machines which was originally developed for classification has also been extended for regression analysis (as further detailed in Chapter 5).

Calibration approaches can be divided into two distinct classes, namely univariate and multivariate. Historically, univariate approaches were used to calibrate a single measurement variable, such as height of a characteristic peak, to the property of interest, such as analyte concentration. However, this approach only takes advantage of a small portion of the data available and is incapable of exploiting the biochemical specificity of Raman spectroscopy. Further, in complex biological samples such as the blood tissue matrix, where several distinct moieties may contribute to the intensity of a specific Raman feature, univariate approaches prove to be inadequate. In contrast, multivariate calibration schemes are able to exploit the multi-channel nature of the spectroscopic datasets to determine concentration information at trace levels in complex mixture samples.

For the simplest form of multivariate calibration (i.e. linear), we assume that the measured spectrum, **s**, can be mathematically expressed as a linear weighted combination of the pure component spectra of the sample constituents, **p**, with the weights being equal to the corresponding analyte concentrations, c:

$$\mathbf{s} = c_1 \cdot \mathbf{p}_1 + c_2 \cdot \mathbf{p}_2 + \cdots + c_n \mathbf{p}_n + \mathbf{e} \qquad (3.9)$$

where the subscripts on the RHS of the equation denote the specific constituents, e.g. glucose, creatinine, urea, cholesterol, keratin etc. for the blood tissue matrix and **e** is the measurement error including noise. (In our discussions pertaining to the multivariate calibration schemes, we will employ the following notation: lowercase boldface to denote a column vector; uppercase boldface to denote a matrix; lowercase body type to denote a scalar quantity; the superscript ^{T} to denote a transpose operation.)

Equation (3.9) presents a reasonably good first approximation based on the incoherent addition of the constituent signals in spontaneous Raman spectroscopy. The linearity of the mixture spectra was also experimentally demonstrated in biological tissue under controlled conditions by Manoharan *et al.* [31].

Extending Eq. (3.9) to account for multiple mixture spectra we get:

$$\mathbf{S} = \mathbf{CP} + \mathbf{E}, \tag{3.10}$$

where **S** is a (j x λ) matrix of sample spectra with each sample spectrum occupying a row, **C** is a (j x p) matrix of constituent concentrations in the samples, **P** is a (p x λ) matrix of constituent spectra with each constituent spectra occupying a row and **E** is the associated measurement error. Here, p is the number of the sample constituents and λ is the number of the discrete spectral information components, e.g., pixels in the detector.

If we neglect the error, Eq. (3.10) can then be re-written in a form suitable to the determination of concentration information from the measured spectral dataset:

$$\mathbf{C} = \mathbf{S}\mathbf{P}^{\mathrm{T}}\left(\mathbf{P}\mathbf{P}^{\mathrm{T}}\right)^{-1} \tag{3.11}$$

Depending on the application, or more precisely on the *a priori* information available for the specific application, Eq. (3.11) can then be employed in different ways. In general, these different multivariate calibration strategies can be divided into two classes, namely explicit and implicit calibration.

Explicit calibration

Explicit multivariate calibration methods necessitate knowledge of the full set of the individual component spectra, either through measurement or appropriate calculation. For example, if all

the component spectra (i.e. spectra of a unit concentration of each pure constituent) are available, then Eq. (3.11) can directly be used to predict the concentration of each analyte. This is the simplest case, which does not involve a 'true' calibration step, since the model **P** is already available. Such a solution is aptly called ordinary least squares (OLS) analysis.

A similar scheme, called classical least squares (CLS) analysis, is employed when the full set of pure component spectra are not known *a priori* but can be calculated by use of an appropriate training or calibration set where the concentrations of all the pure components are known. Under such circumstances we can estimate $\hat{\mathbf{P}}$, where the hat indicates that the model is estimated and is not directly available:

$$\hat{\mathbf{P}} = (\mathbf{C}_{calib}{}^{T}\mathbf{C}_{calib})^{-1}\mathbf{C}_{calib}{}^{T}\mathbf{S}_{calib} \qquad (3.12)$$

where the subscript 'calib' indicates that the model is estimated based on the known concentrations and measured spectra from the calibration dataset. $\hat{\mathbf{P}}$, calculated from Eq. (3.12) can be substituted in place of **P** in Eq. (3.11) to obtain the concentrations of the components in the prediction samples based on the spectra acquired from those samples (i.e. **S**).

The noteworthy aspect of the CLS scheme is that it involves a distinct calibration and prediction step, namely Eq. (3.12) and Eq. (3.11) respectively. Unfortunately, in a large number of applications of biomedical spectroscopy including the blood glucose monitoring problem, neither the full set of component spectra or the concentration information is known. In such cases, one has to employ implicit calibration methods.

Implicit calibration

Implicit calibration methods are primarily applied when there is a single analyte of interest, e.g. in non-invasive blood glucose monitoring. The only information available is the concentrations

of the analyte of interest in the calibration samples and the spectra acquired from them. In other words, there exists no further direct knowledge or interest in the identities and concentrations of the other components of the sample. The objective of the implicit calibration methods can then be mathematically written as:

$$\mathbf{c}_{pred} = \mathbf{s}_{pred}{}^{T}\mathbf{b}, \tag{3.13}$$

where **b** is the spectrum of regression coefficients which when used in conjunction with the spectrum of the prospective sample **s** can provide the analyte concentration, c, in that sample. This equation is analogous to Eq. (3.11) of explicit calibration, where the generalized pseudo-inverse of **P** has been replaced with the so-called b-vector (**b**) for the analyte of interest.

The critical question then is: how do we find the optimal b-vector? Extending Eq. (3.13) for a number of calibration samples, we get:

$$\mathbf{c}_{calib} = \mathbf{S}_{calib}{}^{T}\mathbf{b}, \tag{3.14}$$

where $\mathbf{c}_{calib}$ is the vector of the analyte concentrations in the calibration samples and $\mathbf{S}_{calib}$ is the matrix of measured calibration spectra. In principle, the b-vector for the analyte of interest can be obtained by simply inverting Eq. (3.14). Unfortunately, in practice, the inversion of the matrix $\mathbf{S}_{calib}$ is complicated by the fact that the number of wavelengths is typically much larger than the number of samples. In other words, the expanded matrix ($\mathbf{S}_{calib}{}^{T}\,\mathbf{S}_{calib}$) has inherent degeneracy, which must be removed before the pseudo-inverse can be computed. This means that we must somehow compress the spectral data into fewer data points, or have more samples than the number of data points. The simplest method is to choose a subset of the total number of wavelengths, such that the number of spectral points employed is less than or equal to the number of calibration samples. Unfortunately, this may mean loss of significant spectral

information. The other option is to compress the spectral data into a significantly smaller number of key factors or, principal components. Implicit calibration methods, such as principal component regression (PCR) and partial least squares (PLS), provide the necessary means for such a transformation, also known as reduction in dimensionality of the dataset. Figure 3.5 provides an overview of an implicit calibration procedure specialized for spectroscopy-based blood glucose monitoring.

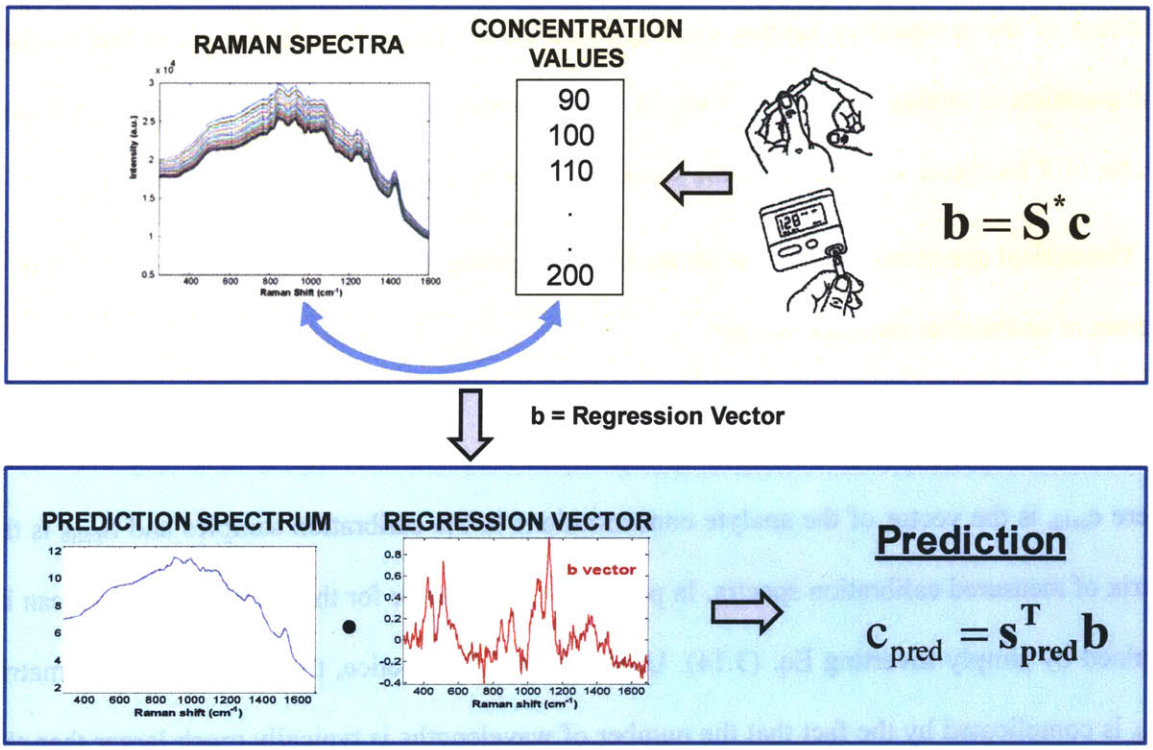


Fig. 3.5 Schematic showing the essential two-step process involved in implicit calibration. Here, c and S represent the concentration and spectra collected from the calibration samples; S* indicates the generalized inverse of S; b is the regression vector; and c_{pred} is the concentration estimated in the prospective sample based on the acquired spectrum, s_{pred}. Further details of the procedures are provided in the text.

PCR uses principal component analysis (PCA) to compute the principal components and subsequently performs regression on a limited number of principal components. PCA decomposes the matrix of calibration spectra into orthogonal principal components that best capture the variance in the spectral dataset. PCA is a linear, though abstract, mathematical transformation of the data into a new coordinate system (axes) such that the largest variance lies on the first axis and decreases thereafter for each successive axis. This has the effect that most of the redundant information and noise can be omitted out of the calibration model if only a few principle components are retained. The critical question then is to decide how many of the principle components should be retained in the subsequent analysis. Typically, different metrics, such as error in leave-one-out cross-validation or cumulative percentage contribution of the retained principle components to the overall variance, are employed to optimize the number of principle components used to construct the calibration model. In an ideal situation, this number should equal the number of sample constituents, i.e. the number of pure components in the mixture. However, for biological applications, there is rarely a situation where such a number is known *a priori*; moreover, correlations between sample constituents, system drift and noise also play a key role in the final number of principle components retained in real world situations [24]. Mathematically, one can view this PCA step as a decomposition of the spectral dataset into scores (**T**) and loading matrices (**Q**):

$$\mathbf{S} = \mathbf{TQ}. \tag{3.15}$$

Here, T is a (j x n) matrix and Q is a (n x λ) matrix, where n is the number of principle components retained for the calibration model. The next step is to the so-called rotation step, which is used to construct a relationship between the concentrations and the scores:

$$\mathbf{c} = \mathbf{Tr}, \tag{3.16}$$

where **r** represents a column vector of size n. Some researchers like to think of **r** as the analogue of the b-vector in the score space, rather than in the spectrum space. To determine **r** from the calibration samples, we simply invert Eq. (3.16):

$$\mathbf{r} = (\mathbf{T}_{calib}{}^{T}\mathbf{T}_{calib})^{-1}\mathbf{T}_{calib}{}^{T}\mathbf{c}_{calib} \tag{3.17}$$

This pseudo-inverse can be readily computed as long as $n \leq j$. Eq. (3.16) for calibration can then be combined with Eq. (3.15) to obtain the final concentration estimate for prediction. In the application of Eq. (3.15) for prediction, the spectrum measured from the prospective sample is decomposed into the appropriate scores vector.

PLS is the extensively used implicit calibration technique. Its formulation is not too dissimilar to that of the PCR formulation presented above, with the key exception that the initial matrix decomposition for PLS is implemented on the covariance matrix of the spectra and the reference concentrations in contrast to PCR where only the variance in the spectral dataset was considered. In other words, PLS assumes that errors exist in both blocks - the measured spectra as well as the reference concentrations - and both are of importance to the analysis. This is of significance in many biomedical spectroscopy applications, where the spectra measurement are often more reproducible in comparison to the reference concentrations, which may be hampered by sample preparation techniques. In practice, PLS and PCR are shown to have similar performance if noise in the spectral data and errors in the reference concentration measurements are negligible. Otherwise, PLS provides better predictive ability than an equivalent PCR model and is therefore more widely used in various chemometric applications. Nevertheless, it should be noted that the principle components obtained by PCA typically provide significant physical interpretability, a trait not shown by the loading vectors of PLS.

Clearly, the implicit calibration methods provide a powerful tool for spectral analysis in a vast number of cases where only limited information of the analyte of interest is present. Unfortunately, it is susceptible to chance correlations, such as covariations among two or more sample constituents (further explored in Chapter 6) and system drift. This issue has previously been addressed by several investigators, who have proposed hybrid calibration methods by combining features specific to explicit and implicit methods [32, 33]. In particular, our laboratory has proposed the incorporation of prior information, e.g. knowledge of glucose spectrum, as an additional constraint to guide the development of the b-vector [34, 35].

Finally, all the aforementioned methods function under the assumption that a linear regression process is sufficient for extraction of concentration information. In recent years, this assumption has come under question, especially when non-linear effects are induced by external variables such as temperature. In such cases, a non-linear regression scheme such as non-linear PLS, artificial neural networks or support vector regression may be more appropriate. Chapter-5 discusses the introduction of such curved effects and how they may be accounted for using non-linear support vector regression.

Validation and figures of merit

In general, multivariate calibration is effective provided that the range of samples used to develop the model is adequately representative of all prospective samples that may be analyzed by this model. Nevertheless, an important feature of the total calibration process is the performance of appropriate validation for the developed model. This is especially true for implicit calibration methods, which cannot be simply judged by conventional statistical measures such as goodness of fit. As noted by other researchers [36], spurious effects can be incorrectly interpreted as legitimate correlations even when no actual information content is present in the

experimental dataset. Such effects may provide positive results, based on non-analyte specific effects, inspiring false confidence in the measurements. To avoid, such possibilities rigorous statistical validation must be performed, especially for cases where the signal from the analyte of interest is a small fraction of the overall signal.

There are two major classes of validation techniques for multivariate calibration, namely internal validation and external validation. In cases where the total number of available samples is limited, internal validation is undertaken on the calibration dataset. The most popular internal validation strategy is "leave-n-out" cross-validation. For this type of validation, n samples are left out of the calibration dataset and used as validation samples. This step is iterated multiple times to get an average estimate of the validation errors. A special case of internal validation is leave-one-out cross-validation, where one data point is left out at a time from the calibration data and the developed model is used to compute the concentration of the left out data point. This procedure is then repeated until all data points have been left out in turn.

It is common to use errors in the concentration block for definitive decisions on the quality of the calibration model. Mathematically, the error in cross-validation, i.e. the root mean square error of cross validation (RMSECV), can be written in the following manner:

$$RMSECV = \sqrt{\frac{\sum_{i}^{n}\left(c_{extracted,i} - c_{reference,i}\right)^2}{n_C - p}}, \qquad (3.18)$$

where $c_{extracted}$ and $c_{reference}$ are the calculated (using the implicit calibration method of choice) and reference concentrations of the validation samples, n_c is the total number of calibration samples and p is the number of components (e.g. principal components and loading vectors) used to construct the calibration model.

Typically, errors in cross-validation will initially decrease, then level off and finally show a slight increase as the number of components used in the calibration model are increased. This suggests that there exists an optimal number of components, which explains most of the information content and is suitable for prospective application. Further addition of components (often referred to as 'over-fitting' or 'over-training') is likely to reduce chances of success in prospective application as later components may introduce significant noise in the calibration model. A useful rule-of-thumb to reduce overfitting is to have three times as many calibration samples as the rank of the calibration model (i.e. the number of principle components) [37].

External validation is similar to internal validation, except that a certain number of samples are withheld from the calibration process for optimization of the calibration parameters. This dataset is often referred to as the tuning dataset. Invoking an independent dataset ensures that the model has not been skewed by data specific to the calibration set. (Readers may also be interested in investigating the bootstrap method, which represents a compromise between the internal and external validation procedures [38].)

Irrespective of the nature of validation, however, the true test lies in prospective application on unknown samples, after the calibration model has been completely finalized. The prediction error, root mean square error of prediction (RMSEP), is computed as:

$$RMSEP = \sqrt{\frac{\sum_{i}^{n}\left(c_{predicted,i} - c_{reference,i}\right)^2}{n_P}}, \qquad (3.19)$$

where n_p is the number of prospective prediction samples.

We conclude our discussion in this section with a cautionary note on the blind application of these multivariate calibration models. The primary challenge in establishing calibration transfer

in prospective samples is the establishment of causation between the analyte (glucose) levels and the optical signals, i.e. whether the source of the signal truly reflects a change in glucose concentrations. For all calibration models (linear or non-linear), spurious effects may arise from different sources that correlate with changes in glucose, especially in human subject studies. The burden of responsibility, then, rests on the investigators to find alternate means of demonstrating that the predicted results are not a direct outcome of such spurious effects. One way of accomplishing this is to use tissue-simulating phantoms under controlled environmental conditions, where *a priori* knowledge of each system component exists. Evidently, performance under such conditions can only be used to investigate specific issues. Second, one may feed false concentration profiles into the calibration scheme to test its capability of prospective prediction. Ideally, such concentration profiles should not be predicted well as they are not reflective of anything in the experimental dataset. Nevertheless, time-dependent profiles often produce spurious results because of correlations with system drift. Such problems have confounded investigators in NIR absorption spectroscopy for decades. While we anticipate that the enhanced specificity of Raman spectroscopy is likely to reduce some of the aforementioned issues, the statistical validation of such techniques by introducing perturbations to the input information should be rigorously pursued.

3.4 Previous research in the MIT Spectroscopy Laboratory

Our laboratory has pioneered the application of NIR Raman spectroscopy to this important diagnostic problem of non-invasive blood glucose monitoring. Over the past decade and a half, novel instrumentation and methodologies have been developed to overcome the myriad technical challenges that may potentially impede the path to successful prospective application in human subjects. In the following, we outline our laboratory's prior *in vitro* and *in vivo* studies to

understand the status of the project before the work presented in this thesis was undertaken and to identify the key challenges based on the results of these previous studies.

3.4.1 *In vitro* studies

In the laboratory's initial studies, quantitative concentration measurements of multiple analytes were demonstrated first in simple mixture samples such as in water solutions. Subsequently, the dataset complexity was increased to include human blood serum and whole blood samples as well as other tissue-simulating phantoms.

Our initial demonstration of concentration measurements in biological media reported measurement of multiple analytes including glucose, urea, cholesterol, triglyceride, albumin, total protein and hematocrit in serum and whole blood samples from sixty-nine patients over a seven-week period [39]. The serum measurement errors were within clinical accuracy requirements, but it was observed that the whole blood measurement errors were considerably higher. The data indicated that the reduced signal levels obtained from whole blood samples were the principal source of increased error in the study. In addition, the laboratory also made two important advances during the period of these *in vitro* studies: (a) development of hybrid linear analysis (HLA), a novel multivariate calibration method that incorporated prior information content (such as spectral line shape) into the implicit calibration method [34]; and (b) formulation of an analytical method for estimating the chemometric prediction error, i.e. the prediction uncertainty [40]. The prediction uncertainty is instrumental in defining the minimum detectible concentration of the system.

To overcome the problem of reduced signal collection in the investigations of whole blood samples, a new system incorporating non-imaging optical elements was assembled. This was

found to improve the signal collection capability by a factor of >4. A follow-up whole blood study using this improved instrument demonstrated the feasibility of measuring multiple analytes in whole blood with clinical accuracy [15]. In this study, whole blood samples were collected from 31 patients undergoing routine clinical evaluation. 30 consecutive 10-second spectra were acquired from each such sample. Conventional clinical laboratory methods were used to measure eight reference analyte concentrations, including glucose, urea, total protein, albumin, triglycerides, hematocrit and hemoglobin. These reference concentrations were correlated with the recorded Raman spectra through multivariate calibration and validation. The results demonstrated that the Raman spectroscopy based concentration predictions had a very strong correlation with the reference concentrations (the square of correlation coefficient between the reference and predicted concentrations was greater than or equal to 0.93) for all analytes, except total cholesterol.

3.4.2 *In vivo* studies

The promising results of these *in vitro* studies provided the impetus for an initial *in vivo* evaluation of the ability of Raman spectroscopy to measure glucose in a transcutaneous manner [29]. The *in vivo* studies were conducted in a small cohort of human subject volunteers. In the human subject study, a series of Raman spectra were collected from the forearms of healthy human volunteers in conjunction with an oral glucose tolerance test. A glucose tolerance test is a standard diabetes screening procedure, where a subject is typically given a glucose-rich drink (for the oral version of the test) in order to induce a substantive rise in the subject's blood glucose levels. Blood samples are then withdrawn at specific time intervals to investigate the rate of clearance of glucose from the blood stream and therefore to infer the effectiveness of the subject's insulin-based glucose regulation mechanism.

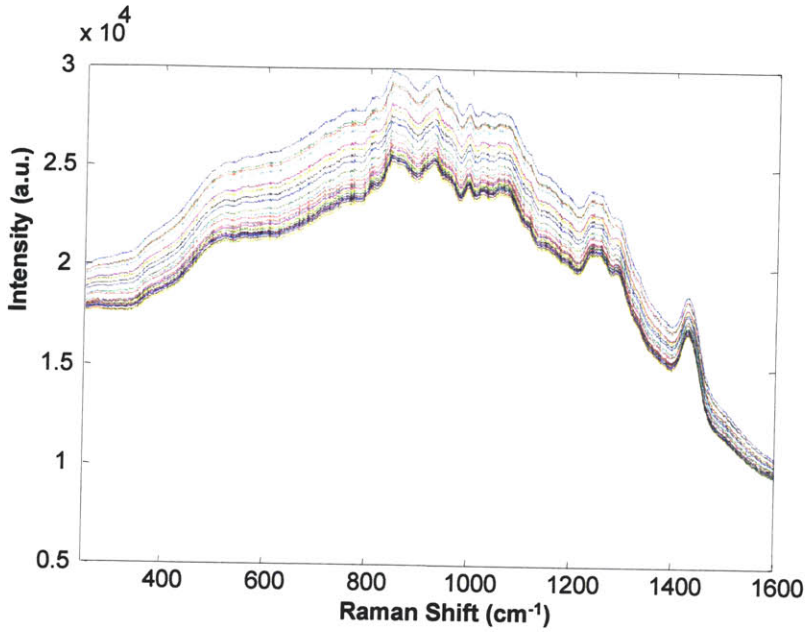


Fig. 3.6 Representative Raman spectra acquired from a human volunteer during an oral glucose tolerance test study.

Raman spectra and reference blood glucose concentrations were measured at regular 10 minute intervals during the 2-3 hour duration of test. Figure 3.6 shows a set of representative spectra acquired from one such study. A Hemocue® glucose analyzer was employed to obtain the reference blood glucose measurement. A calibration model was generated *individually* from the data from each volunteer using PLS with leave-one-out cross validation. The RMSECV value for the whole dataset based on individual calibration was computed to be 18 mg/dL with the coefficient of determination (square of the sample correlation coefficient) equal to 0.87. Although single-subject cross validation resulted in acceptable error estimates, prospective prediction on multiple subjects proved more difficult. At this time, similar errors were also anticipated for prospective prediction on the same human subject at future times. In other words, the reproducibility of the developed calibration time was considered to be an open question.

Given the results of this initial *in vivo* investigation, a more methodological implementation and validation from bench to bedside was deemed to be necessary in order to make non-invasive glucose measurement a viable clinical technology. These investigations were meant to specifically address the root causes of the problem, i.e. the factors that cause the substantive increase in prediction error in prospective application.

In this context, Shih and co-workers initiated research along two major directions: (a) development of a flexible hybrid calibration model, called constrained regularization, which was reported to provide some key advantages over partial least squares and hybrid linear analysis [35]; and (b) improvements in quantitative Raman spectroscopy based predictions in turbid media where sample-to-sample variations in sampling volume may otherwise cause significant errors [41, 42].

This thesis significantly extends this initial work in identifying and accounting for the specific technical challenges that impede the path to true prospective application and thereby to clinical translation. Application of the new methodologies developed in this thesis will not only help from a clinical accuracy standpoint but also provide greater insight into the fundamental physiology and physics of such measurements. Finally, the sequential increase in data complexity, including the use of suitable tissue-simulating phantoms at each step, prevents the potential incorporation of spurious and misleading effects that may otherwise creep in, given the large range of uncontrolled variables in clinical studies.

3.5 References

(1) Gardiner, D. J.; Graves, P. R. *Practical Raman Spectroscopy* Springer, **1989**.

(2) Smekal, A. *Naturwiss* P873, Academic Press, **1923**.

(3) Raman, C. V.; Krishnan, K. S. *Nature* **1928**, 122, 12.

(4) Ozaki, Y. *Applied Spectroscopy Reviews* **1988**, 24(3), 259-312.

(5) Hanlon, E. B.; Manoharan, R.; Koo, T. W.; Shafer, K. E.; Motz, J. T.; Fitzmaurice, M.; Kramer, J. R.; Itzkan, I; Dasari, R. R. Feld, M. S. *Physics in Medicine and Biology* **2000**, 45, R1-R59.

(6) Mahadevan-Jansen, A.; Richards-Kortum, R.; Engineering in Medicine and Biology Society, Proceedings of the 19th Annual International Conference of the IEEE , **1997**, 6, 2722-2728.

(7) Haka, A. S.; Shafer-Peltier, K. E.; Fitzmaurice, M.; Crowe, J.; Dasari, R. R.; Feld, M. S. *Proceedings of National Academy of Sciences* **2005**, 102(35), 12371-12376

(8) Long, D.A. *The Raman Effect* John Wiley & Sons, New York, **2002**.

(9) Weber, A. *Raman Spectroscopy of Gases and Liquids* Topics in Current Physics, Springer-Verlag, Berlin, **1979**.

(10) McCreery, R. L. *Raman Spectroscopy for Chemical Analysis* John Wiley & Sons, New York, **2000**.

(11) Szymanski, H. A. *Raman Spectroscopy, Theory and Practice Vol. 1 and 2* Plenum, New York, **1967**.

(12) Schrader, B. *Infrared and Raman Spectroscopy* John Wiley & Sons, New York, **1995**.

(13) Tuchin, V. *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis, Second Edition* SPIE Publications; 2nd edition, **2007**.

(14) Hirschfeld, T.; Chase, B. *Applied Spectroscopy* **1986**, 40, 133-137.

(15) Enejder, A. M. K.; Koo, T. W.; Oh, J.; Hunter, M.; Sasic, S.; Feld, M. S.; Horowitz, G. L. *Optics Letters* **2002**, 27, 2004-2006.

(16) K. Tanaka, M. T. T. Pacheco, J. F. Brennan III, I. Itzkan, A. J. Berger, R. R. Dasari and M. S. Feld, *Applied Optics*, 1996, **35**, 758-763.

(17) Matousek, P. *Chemical Society Reviews* **2007**, 36, 1292–1304.

(18) Barman, I.; Tan, K. M.; Singh, G. P. *Journal of Raman Spectroscopy* **2010**, 41(10), 1099-1101.

(19) Motz, J. T.; Fitzmaurice, M.; Miller, A.; Gandhi, S. J.; Haka, A. S.; Galindo, L. H.; Dasari, R. R.; Kramer, J. R.; Feld, M. S. *Journal of Biomedical Optics* **2006**, 11, 021003.

(20) Shih, W. C. *Quantitative Biological Raman Spectroscopy for Non-invasive Blood Analysis*, MIT Ph.D. Thesis, **2007**.

(21) Soderholm, S.; Roos, Y. H.; Meinander, N.; Hotokka, M. *Journal of Raman Spectroscopy* **1999**, 30(11), 1009-1018.

(22) Bell, A. F.; Barron, L. D.; Hecht, L. *Carbohydrate Research* **1994**, 257(1), 11-24.

(23) Arnold, M. A.; Bechtel, K. L.; Shih, W. C.; Feld, M. S. *Multivariate calibration. In Analytical Chemistry of in vivo Glucose Measurements J. A. Stenken and D. D. Cunningham (eds)*, Wiley, **2008**.

(24) H. Martens and T. Naes, *Multivariate Calibration*, Wiley, New York, **1989**.

(25) Brereton, R. G. *Chemometrics : data analysis for the laboratory and chemical plant*; Wiley: Chichester, West Sussex, England ; Hoboken, NJ, **2003**.

(26) Massart, D. L. *Handbook of chemometrics and qualimetrics*; Elsevier: Amsterdam ; New York, **1997**.

(27) Reddy, R. K.; Bhargava, R. *Chemometric Methods for Biomedical Raman Spectroscopy and Imaging* in *Emerging Raman Applications and Techniques in Biomedical and Pharmaceutical Fields, Matousek, P. and Morris, M. D.(eds)*, SpringerLink, **2010**.

(28) Katsumoto, Y.; Ozaki, Y. *Applied Spectroscopy* **2003**, 57(3), 317–322.

(29) Enejder, A. M. K.; Scecina, T. G.; Oh, J.; Hunter, M.; Shih, W.-C.; Sasic, S.; Horowitz, G.; Feld, M. S. *Journal of Biomedical Optics* **2005**, 10, 031114.

(30) Savitzky, A.; Golay, M. J. E. *Analytical Chemistry* **1964**, 36 (8), 1627–1639.

(31) Manoharan, R.; Shafer, K.; Perelman, L.; Wu, J.; Chen, K.; Deinum, G.; Fitzmaurice, M.; Myles, J.; Crowe, J.; Dasari, R. R.; Feld, M. S. *Photochemistry and Photobiology* **1998**, 67, 15-22.

(32) Wentzell, P. D.; Andrews, D. T.; Kowalski, B. R. *Analytical Chemistry* **1997**, 69, 2299-2311.

(33) Haaland, D. M.; Melgaard, D. K. *Applied Spectroscopy* **2001**, *55*, 1-8.

(34) Berger, A. J.; Koo, T.W.; Itzkan, I.; Feld, M. S. *Analytical Chemistry* **1998**, 70(3), 623-627.

(35) Shih, W. C.; Bechtel, K. L.; Feld, M. S. *Analytical Chemistry* **2007**, 79 (1), 234-239.

(36) Arnold, M. A.; Burmeister, J. J.; Small, G. W. *Analytical Chemistry* **1998**, 70, 1773-1781.

(37) Qi, D.; Berger, A. J. Applied Optics **2007**, 46, 1726-1734.

(38) Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*, Chapman & Hall, New York, **1993**.

(39) Berger, A. J.; Koo, T.W.; Itzkan, I.; Horowitz, G.; Feld, M. S. *Applied Optics* **1999**, 38(13), 2916-2926.

(40) Berger, A. J.; Feld, M. S. *Applied Spectroscopy* **1997**, 51(5), 725-732.

(41) Shih, W. C.; Bechtel, K. L.; Feld, M. S. *Optics Express* **2008**, 16, 12726-12736.

(42) Bechtel, K. L.; Shih, W. C.; Feld, M. S. *Optics Express* **2008**, 16, 12737-12745.

CHAPTER 4

TURBIDITY-CORRECTED RAMAN SPECTROSCOPY

A major challenge in quantitative biological Raman spectroscopy, particularly as applied to non-invasive measurements on biological tissue, lies in overcoming the detrimental effects of tissue scattering and absorption. In this work, the interplay of sample (elastic) scattering and absorption is referred to as tissue turbidity. The Raman spectral information is corrupted by multiple scattering and absorption events in the surrounding medium, causing non-analyte specific variances and thereby diminishing the predictive capability of the calibration model. To account for these distortions in the intrinsic Raman spectra, a novel analytical method, Turbidity Corrected Raman Spectroscopy (TCRS), is proposed in this chapter. This method is based on the photon migration approach and employs alternate acquisition of diffuse reflectance and Raman spectra to compensate for the effects of turbidity in the latter. It is demonstrated that, on application of TCRS, the widely varying Raman spectra observed from a set of tissue phantoms having the same concentration of Raman scatterers but different turbidities tend to coalesce onto a single spectral profile. Furthermore, in a prospective study employing physical tissue models with varying turbidities and randomized concentrations of Raman scatterers and interferents, an average reduction of 20% in prediction error is obtained by applying the turbidity correction procedure to the acquired Raman spectra.

4.1 Significance

As mentioned in Chapter 3, several research groups, including our own, have reported Raman spectroscopy based glucose predictions at physiologically relevant concentrations in serum [1],

whole blood [2] and other *in vitro* samples, such as human eye aqueous humor [3, 4]. Promising studies in transcutaneous Raman spectroscopy in human volunteers have also been conducted [5, 6]. However, prospective application has proven to be challenging, particularly when the calibration model developed on one subject is applied to a different subject. A significant portion of the difficulty in implementing successful prospective prediction methods can be attributed to variations in tissue optical properties characteristic of a human subject population. Variations in turbidity, defined as the combined effects of scattering and absorption, alter the tissue sampling volume thereby diminishing the concentration prediction accuracy on prospective samples. In addition, tissue turbidity also gives rise to intensity and shape distortions in the acquired Raman spectra and introduces non-analyte specific variances into the calibration model that deteriorates the prediction capability. Overcoming the effects of optical property variations is one of the central challenges in quantitative biological Raman spectroscopy. In other words, there is a clear need for a methodology that enables the extraction of intrinsic (true) line shape and intensity information from the acquired Raman spectra in order to accurately predict blood analyte concentrations in a group of human subjects.

One approach developed to reduce the aforementioned effects of the variations in optical properties on the prediction capability of the regression model was presented by Chaiken *et. al.* [7]. They conducted an *in vivo* study employing mechanical tissue modulation on the fingertip and the difference of the spectra acquired at non-modulated and modulated states, respectively, to enhance the spectral contributions from the blood pool. Although reasonable glucose measurement accuracy was reported, the results showed that differences among individuals added significant error to the actual concentration estimates.

The need for correction of turbidity induced variations in sampling volume and the intrinsic spectral shape has been widely acknowledged in the non-invasive glucose detection literature, especially near-infrared (NIR) absorption spectroscopy [8-11]. Over the past decade, researchers in the field of fluorescence spectroscopy have also extensively studied the effects of turbid media on the observed fluorescence spectra [12, 13]. Correction schemes using a variety of experimental and analytical approaches have been developed to remove distortions induced in the fluorescence spectra due to multiple scattering and absorption [14-22]. For example, our laboratory has developed intrinsic fluorescence extraction methodologies using diffuse reflectance spectroscopy, based on the principle that fluorescent and diffusely reflected photons undergo similar scattering and absorption events in the turbid medium [14, 15].

Even though methods for extracting intrinsic fluorescence have been widely employed, relatively less work has been done in developing an analogous methodology for Raman spectroscopy. In one such scheme employed to study powder samples, the Raman signal was related to the measured diffuse reflectance as a function of the Kubelka-Munk absorption or scattering coefficients [23]. However, the Kubelka-Munk formulation, which assumes that elastic scattering is isotropic, cannot be readily extended to biological tissues, where scattering is highly anisotropic (typically heavily forward-scattered) [24, 25]. Aarnoutse *et. al.* [26] have also reported a method using simultaneous absorption measurements to correct for the absorbance of Raman scattered light by the media in catalysis reactions. In this method, only the effect of variations in sample absorption was considered, whereas in biological tissues, variations in scattering are substantially more important [27, 28]. Matousek *et. al.* have proposed a novel method (spatially offset Raman spectroscopy, SORS) for retrieval of Raman spectra from subsurface layers in diffusely scattering media, which is based on the idea of collecting Raman

scattered light from surface regions laterally offset away from the excitation laser spot [29]. This method, however, is useful for extracting depth-resolved spectral information and is not suitable for application in samples where Raman scatterers are distributed throughout the turbid medium of interest.

Previously, our laboratory has developed a method to reduce the deleterious effects of turbidity variations in concentration predictions using Raman spectroscopy. As in the aforementioned case of fluorescence correction, this method is based on the use of diffuse reflectance at the same tissue region to correct for turbidity variations. Two of our laboratory's previous publications have presented a method (Intrinsic Raman Spectroscopy (IRS)) that corrects for the turbidity-induced sampling volume variations and relates the observed (sampled) concentration to the true concentration of the analyte of interest by means of a universal calibration curve [30, 31]. However, using a Monte Carlo technique (as proposed in IRS [30]) to determine a universal calibration curve has the limitation that the Raman scattering coefficients must be known accurately for the given constituents, which is a non-trivial task. In addition, for the experimental IRS technique [31] used for derivation of the calibration curve, one needs to create tissue phantoms with the same composition of constituents as found in each tissue type. Since glucose forms a minute part of the total tissue Raman spectrum (as small as ~0.3%), there remains substantial uncertainty as to the prospective applicability of this calibration curve *in vivo*. Furthermore, it also requires the explicit determination of the optical properties of the tissue site, specifically the reduced scattering coefficient ($\mu_s'(\lambda)$). Prediction of blood analyte concentrations from human subjects under clinical conditions necessitates the introduction of a new method which is independent of tissue composition and geometry invariant.

In the following, we present an alternate approach, which we call turbidity corrected Raman spectroscopy (TCRS), that overcomes these difficulties by correcting for intensity and shape distortions in the observed Raman spectra. This approach also employs alternate acquisition of diffuse reflectance spectra and is based on the photon migration picture. In the following, we derive an analytical relation connecting the observed Raman spectrum, the diffuse reflectance spectrum and the turbidity corrected Raman spectrum. We then study a set of tissue phantoms having the same concentration of Raman scatterers, but different background turbidities, and demonstrate that on application of TCRS, the turbidity-induced distortions in intensity and line shapes are removed, and the spectra tend to collapse onto a single spectral profile. The results show that TCRS can recover intrinsic line shape and intensity information from the acquired Raman spectrum in a turbid medium for the range of optical properties found in biological tissues. Furthermore, in a set of prospective studies which employ phantoms of varying turbidities and randomized concentrations of Raman scatterers and interferents, we obtain a 20% reduction in prediction error by applying the turbidity correction procedure to the observed Raman spectra. To the best of our knowledge, this is the first demonstration of enhanced prospective prediction accuracy employing turbidity corrected Raman spectra, and it opens the pathway to improved prospective prediction accuracy in transcutaneous Raman spectroscopy.

4.2 Theoretical formulation

Turbidity corrected Raman spectroscopy employs an analytical model that expresses the turbidity corrected Raman spectrum in terms of the observed Raman and diffuse reflectance spectra. Our mathematical formulation builds on the probabilistic framework of the photon migration approach. The photon migration approach, first introduced by Weiss and co-workers [32, 33], has been extensively used to quantify the diffuse reflectance and fluorescence spectra

acquired from biological tissues by other investigators (notably [14] and [15]) (Fig. 4.1). Other approaches used to commonly describe light propagation in turbid media include the radiative transfer equation (and the associated diffusion approximation formulation) and the numerical approaches, such as Monte Carlo simulations. The photon migration approach presents a compromise between these purely analytical and numerical formulations in terms of its accuracy over a wide range of system and sample parameters (e.g. small source-detector separation) and ease of physical interpretability. Further discussion and comparison of these different approaches towards modeling of light propagation in turbid media is outside the scope of this thesis; the reader is referred to excellent references on this subject [34, 35].

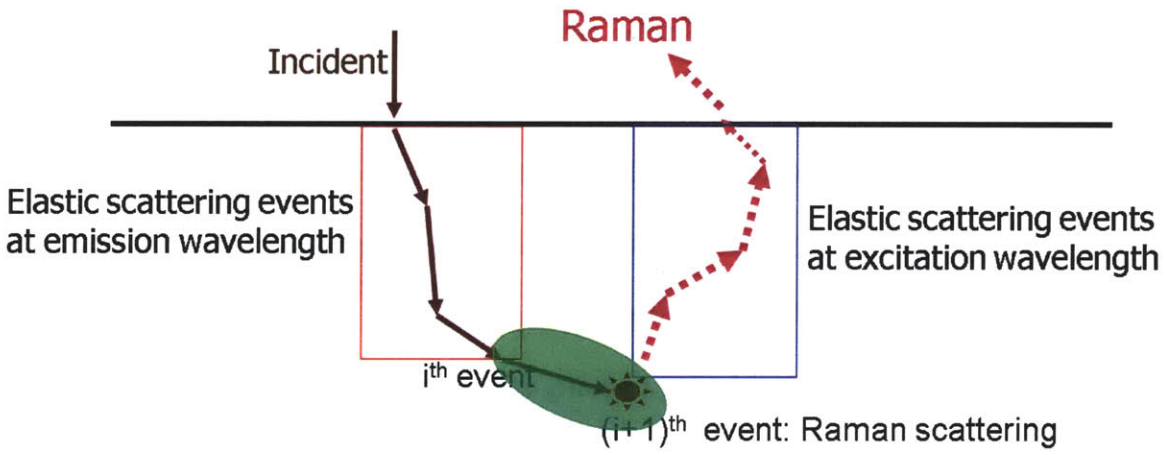


Fig. 4.1 Schematic representation of the photon-tissue interactions for a Raman scattered photon

In the photon migration approach, light propagation in turbid medium is modeled as the ensemble average of the probabilistic photon migration paths distributed within the turbid medium and the realization probability of the full path of each individual photon, governed by the albedo $a = \mu_s/(\mu_s+\mu_a)$, with μ_s and μ_a the scattering and absorption coefficients, respectively. Mathematically, the diffuse reflectance spectrum, R, can then be expressed as:

$$R = \sum_{n=1}^{\infty} \rho_n a^n, \tag{4.1}$$

where ρ is the probability of a photon escaping the input face of the medium after n tissue interaction events and is related to the photon migration path through the medium. The escape probability function is characterized by the scattering phase function and the excitation-collection geometry of the system:

$$\rho_n = S(1-g)\exp[-S(1-g)n], \tag{4.2}$$

where g is the anisotropy coefficient and S is an instrument-specific constant that depends on the excitation-collection geometry [15]. The instrument parameter S must be calibrated for the system under consideration before further spectral analysis can be carried out.

Using the correspondence between Raman scattered and diffusely reflected photons, the observed Raman spectrum can be defined as the ensemble average of the photon migration paths and the realization probability of the individual paths, which is now mediated not only by the absorption and elastic scattering coefficients but also by the relevant Raman scattering coefficient (arising from the analyte concentration and the differential Raman scattering cross-section):

$$(Raman_{OBS})_{xm} = \sum_{n=1}^{\infty}\left[\frac{I_x}{h\nu_x}\sum_{i=0}^{n-1}(h\nu_m)\rho_{ni}\left[a_x^i\left(\frac{\mu_{Rx}}{\mu_{ax}+\mu_{sx}}\right)a_m^{n-i-1}\right]\right], \tag{4.3}$$

where the subscripts x and m represent the Raman excitation (ν_x) and emission (ν_m) frequencies, respectively; $Raman_{OBS}$ is the observed Raman spectrum; h is Planck's constant; i is the position of the node at which Raman scattering occurs; I_x is the excitation intensity; and μ_R is the Raman scattering coefficient.

From Eq. (4.3), it can be observed that an incident photon at frequency v_x undergoes i scattering and absorption events prior to undergoing an inelastic Raman scattering event, with probability $\mu_{Rx}/(\mu_{ax} + \mu_{sx})$, and then undergoes ($n$-$i$-1) further scattering and absorption events at frequency v_m before escaping from the medium. It is worth noting that the escape probability function, ρ_{ni}, is also influenced by the position of the node at which Raman scattering occurs because the scattering coefficient, before and after Raman scattering, has different values at the excitation and emission frequencies [15]. Importantly, the role of photon path in the Raman scattering process can be clearly described from Eq. (4.3), and the similarity between diffuse reflectance and Raman photon paths should be noted.

The turbidity-corrected Raman spectrum, which we define as the spectrum that would be obtained if Raman scattering was the only photon interaction event in the medium, is given by:

$$(Raman_{TC})_{xm} = \frac{I_x}{h\nu_x}(\mu_{Rx} l)h\nu_m \ , \tag{4.4}$$

where $Raman_{TC}$ is the turbidity-corrected Raman spectrum; and l is the effective photon path length in the medium. Specifically, the effective photon path length in turbid tissue is a function of the absorption and scattering properties of the sample.

Relating Eq. (4.4) to Eq. (4.3) using the diffuse reflectance defined in Eq. (4.1), we obtain:

$$(Raman_{TC})_{xm} = \frac{(Raman_{OBS})_{xm}}{\frac{(R_{0m}R_{0x})^{1/2}}{S\mu_{sx}' l}\frac{R_x}{R_{0x}}\left(\frac{R_m}{R_{0m}} + S(1-g)\right)}, \tag{4.5}$$

where R_0 is the diffuse reflectance spectrum that would be obtained from a sample with the same scattering coefficient but zero absorption coefficient, and μ_s' is the reduced scattering coefficient ($\mu_s'=\mu_s(1-g)$).

When the absorption in the tissue is weak (i.e., $\mu_s' >> \mu_a$), Eq. (4.5) can be simplified to obtain:

$$(Raman_{TC})_{xm} = [S][\mu_{sx}' l]\left[\frac{(Raman_{OBS})_{xm}}{\sqrt{R_x R_m}}\right], \tag{4.6}$$

This approximation is valid for biological samples of interest in the NIR range (700-1000 nm), where it is typically observed that $5<\mu_s'<20$ cm^{-1} and $0.08<\mu_a<1.3$ cm^{-1} [36, 37].

Equation (4.6) is key to implementing TCRS and is comprised of three factors: (a) the instrument-specific constant, S; (b) the sample-dependent factor ($\mu_{sx}'l$) (as the optical path length through the tissue is a function of the scattering and absorption properties of the tissue themselves); and (c) the diffuse reflectance-based turbidity correction factor ($(Raman_{OBS})_{xm}/\sqrt{(R_x R_m)}$). As can be seen, this factor depends on the diffuse reflectance measured at both Raman excitation and emission frequencies. This accounts for the fundamental feature of Raman scattering in which a frequency shift from ν_x to ν_m is induced in the outgoing (Raman scattered) photons (as opposed to elastic scattering).

Before applying Eq. (4.6), we need to determine the instrument-specific constant (first factor in Eq. (4.6)) and the dependence of the average path length on the optical properties of the tissue (second factor in Eq. (4.6)). The optical properties can, in turn, be estimated by fitting the observed diffuse reflectance spectra to the model of Zonios *et. al.* [38]. It is important to note that when applying a standard partial least squares (PLS) protocol [39] using the turbidity corrected Raman spectra, there is no need to determine the instrument-specific constant, as it cancels out during the calibration/prediction procedure. Nevertheless this constant is required for performing explicit multivariate calibration studies such as ordinary least squares (OLS) [39]. In

the calibration study of Chapter-4.4, we determine this constant, as well as the functional dependence of the average photon path length on the sample optical properties.

Additionally, it is necessary to estimate the diffusely reflected light at the Raman excitation frequency (R_x), as all frequencies close to the laser line are cut off by the notch filter that is used to suppress the elastically scattered light. Due to the absence of any significant absorbers in biological tissue in the NIR range and the slowly varying nature of the scattering coefficient, the diffusely reflected light at the Raman emission frequency can be readily estimated by extrapolating a low-order polynomial which provides the best fit to the observed diffuse reflectance spectrum at the excitation frequency.

4.3 Experimental Methods

4.3.1 Incorporation of diffuse reflectance

The experimental setup used for implementation of the proposed methodology is similar to that previously described in Chapter-3, except in one crucial aspect - namely, the addition of a broadband source for acquisition of diffuse reflectance spectra. A schematic of this modified experimental arrangement is shown in Fig. 4.2. Briefly, an 830-nm external cavity diode laser (Process Instruments) was used as the Raman excitation source for acquisition of Raman spectra. The laser beam was focused onto the sample with an average power of ~100 mW and a spot diameter of ~1 mm. Additionally, a tungsten-halogen lamp (Avantes AvaLight-HAL-S) was used as a broadband source to obtain diffuse reflectance spectra from the sample. The beam diameter and average power of the broadband source at the sample were measured to be approximately 1 mm and 100 μW, respectively. The tissue phantoms were held in a standard fused silica cuvette. The sample interface was aligned so as to enable the specular reflection,

which does not contain any analyte information, to escape through the hole in the paraboloidal mirror.

The back-scattered light was collected and directed by the paraboloidal mirror towards a notch filter to suppress the excitation (Rayleigh) peak. The light was then focused using an optical fiber bundle onto the slit of a modified f/1.4 spectrograph (Kaiser Optical Systems, Inc.) before final collection by the liquid nitrogen cooled deep depletion CCD detector (1"×1") (Princeton Instruments).

The Raman and diffuse reflectance spectra were acquired alternately with an acquisition time per sample of 20 seconds each. The diffuse reflectance spectra were calibrated using a reference standard of 20% intralipid solution. Spectra of samples were acquired randomly with respect to the constituents' concentrations. Spectra were collected three times over a one week period. The acquired spectra at different times for each phantom were identical within the shot noise variability. In other words, no significant deterioration of the constructed tissue phantoms (arising from reaction between the phantom constituents) was observed over the one week period.

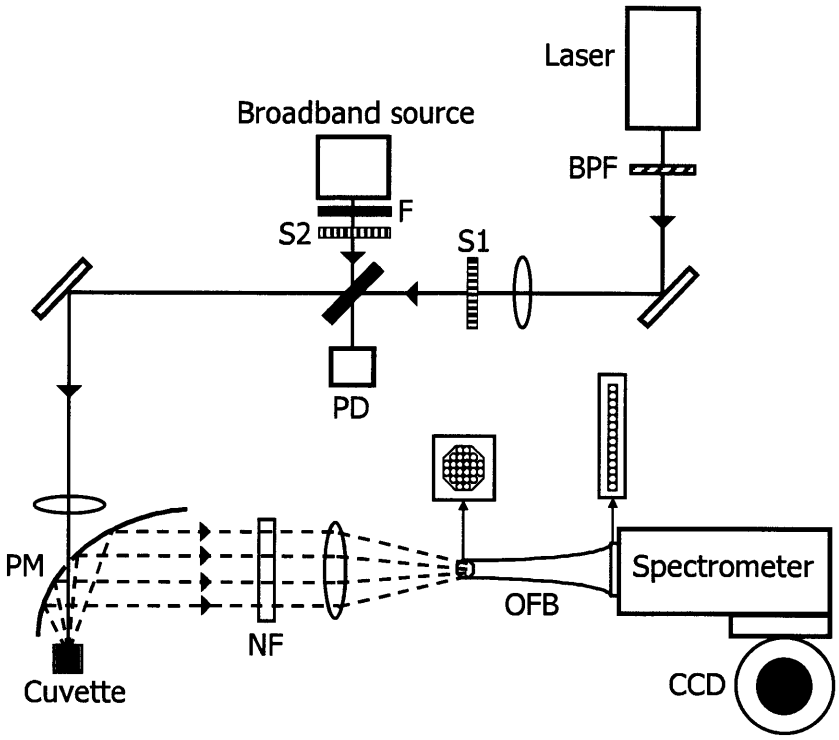


Fig.4.2. Schematic of the modified experimental setup for TCRS implementation. BPF: bandpass filter; S1-S2: shutters; F: Absorption filter; PD: photodiode; PM: paraboloidal mirror; NF: notch filter; OFB: optical fiber bundle.

4.3.2 Physical Tissue Models

Tissue phantom studies were performed to (1) calibrate the instrument; (2) validate the TCRS formalism; and (3) analyze its performance in prospective prediction, respectively. The objective of the calibration study (1) was to determine the functional dependence of the second factor of Eq. (4.6) ($\mu_{sx}'l$) on the optical properties of the samples, and to estimate the instrument-specific constant S (first factor of Eq. (4.6)). The validation study (2) was designed to characterize the effectiveness of the TCRS methodology in correcting turbidity induced variations in the

observed Raman spectra. For both of these studies, ordinary least squares (OLS) regression was employed to find the concentration of the components, especially that of the analyte of interest, as OLS provides the most accurate results when the spectra of all the analytes can be individually measured, as in this case. The spectra of the tissue model constituents, which were used in the OLS analysis, are shown in Fig. 4.3. Employing OLS for the calibration and validation studies ensures that the first and second factors of Eq. (4.6) are determined accurately and that the effectiveness of TCRS is analyzed in the absence of any spurious correlations that might be generated in an implicit calibration strategy.

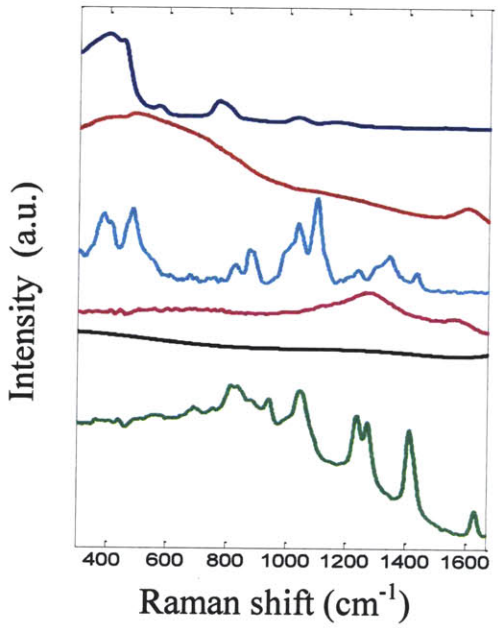


Fig. 4.3. Spectra of OLS model components: cuvette, water, glucose, India ink, fluorescence background and intralipid (shown in order from top to bottom and offset for clarity)

Finally, the prospective prediction study (3) was performed to quantitatively test the prospective prediction capability of TCRS. This study simulated the clinical conditions in which the Raman and diffuse reflectance spectra are being acquired from a human subject. Under such circumstances, in which the spectra (or concentrations) of all constituents are unknown *a priori*,

an implicit calibration framework must be employed. Thus, to replicate the protocol followed in a clinical study, the results of the prospective prediction study were analyzed using partial least squares (PLS) analysis.

The tissue phantoms were prepared by pipetting glucose (Sigma Aldrich), creatinine (Sigma Aldrich), intralipid (Baxter Healthcare), India ink (Super Black India Ink, Speedball Art Products Co.) and distilled water into clean glass vials. Aliquots of these tissue models were taken in a fused silica cuvette for spectroscopic measurements performed at room temperature. Between each measurement, the cuvette was washed three times with distilled water and dried. The phantoms were stored in the refrigerator at 4°C to avoid contamination.

India ink and intralipid were used as the primary absorber and scatterer to simulate tissue absorption and scattering in the NIR range. Specifically, the scattering and absorption coefficients were varied from 24 to 130 cm^{-1} (at 830 nm) and 0.08 to 1.3 cm^{-1} (at 830 nm), respectively, similar to the ranges observed in human biological tissue. The values of the anisotropy parameter (g) used in the tissue models, 0.8-0.9 [40], were chosen to be comparable to the typical values of g observed in biological tissue. Glucose and creatinine, both of which have characteristic Raman spectral profiles, functioned as the analyte of interest and spectral interferent, respectively. All three tissue phantom studies – (1) calibration, (2) validation and (3) prospective prediction – were performed multiple times on different days to examine the reproducibility of the TCRS approach.

In the calibration and validation studies, 20 and 36 tissue phantoms were constructed, respectively. In these studies, the Raman scatterer (glucose) concentration was kept constant at 500mM. The elastic scattering and absorption coefficients were varied randomly in the ranges mentioned above. In both calibration and validation studies, creatinine was not employed, as the

aim was to characterize the effectiveness of TCRS in extracting intrinsic line shape and intensity information of glucose Raman spectrum acquired from a turbid medium.

In the prospective prediction study, 48 tissue phantoms of randomly varying turbidities were prepared with randomized concentrations of glucose and creatinine in the range 4-30mM. This range of glucose concentrations spans hypoglycemic to hyperglycemic levels, in excess of that typically observed in human subjects. The scattering and absorption coefficients were randomly varied over the same ranges used in the calibration and validation studies.

4.4 Results and Discussion

4.4.1 Calibration Study

The observed Raman spectra from the 20 samples, used to calibrate the instrument and to characterize the dependence of the average photon path length on the sample optical properties, showed a significant spread in the spectral profiles over the entire NIR wavelength region. It should be noted that the concentration of glucose in all the samples was kept constant; therefore the Raman spectra under ideal circumstances (i.e. in the absence of multiple scattering and absorption) would be expected to be the same for all the samples. The resulting difference between the observed spectra and the spectra that would be expected in a clear medium can be attributed to variations in turbidity over the tissue phantoms.

To quantify the spread of the observed spectra, we define a "spread index", which is evaluated as the ratio of the difference in maximum and minimum intensity to the maximum intensity at the same specific wavelength for all spectra. The mean spread index calculated over all wavelengths for the observed spectra from 20 tissue phantoms was found to be 7.2%. On application of the diffuse reflectance-based turbidity correction factor ($(Raman_{OBS})_{xm}/\sqrt{(R_x R_m)}$, third factor in Eq. (4.6)), the mean spread index was reduced to 1.62%. In other words, this

reduced spread index was obtained by application of Eq. (4.6), where the first and second factors were held constant. The observed Raman spectra tend to collapse onto a single spectral profile on application of the diffuse reflectance-based turbidity correction factor alone.

It is expected that there would be a negligible dependence of the second factor of Eq. (4.6), $\mu_{sx}'l$, on the tissue optical properties, as the average photon path length is expected to be inversely proportional to the reduced scattering coefficient in weakly absorbing medium. Nevertheless, an optimization procedure was used to estimate the representative function for the average path length (for the second factor in Eq. (4.6)) that minimizes the magnitude of the mean spread index for the 20 tissue phantoms. This minimization is an indicator of better spectral collapse than that obtained by using the diffuse reflectance-based turbidity correction factor alone, thereby enhancing the ability of TCRS in extracting more precise intrinsic line shape and intensity information from the acquired spectra.

A power law function of absorption (μ_a) and reduced scattering coefficients (μ_s') was chosen to represent the dependence of average photon path length on the optical properties of the turbid media. It is to be noted that during this optimization procedure, the instrument dependent constant S (first factor of Eq. (4.6)) was set equal to 1 and the power law function was multiplied with a unit dimensional constant to match the length units of the optical path length, l. Varying the power law exponents, we observed that the minimum spread index was obtained when the average path length was set equal to $(\mu_s')^{-0.97}(\mu_a)^{0.1}$. This value of the minimum spread index was 1.58%, as compared to the 1.62% we obtained using only the diffuse reflectance-based turbidity correction factor in Eq. (4.6). The obtained weak dependence $((\mu_s')^{0.03}(\mu_a)^{0.1})$ of this factor ($\mu_s'l$) on the tissue optical properties is expected, as mentioned before. Due to this weak functional form and the subsequent negligible reduction shown in the mean spread index function, we

conclude that this factor is very weakly dependent on tissue optical parameters. We ignore the dependence of μ_a and μ_s' on this factor in further analysis.

To determine the instrument-specific constant, S, we used the turbidity corrected spectra (obtained after application of the third factor of Eq. (4.6)) in conjunction with the Raman spectra of the tissue phantom constituents in a standard OLS protocol. The instrument-specific constant that yielded the smallest prediction error was S=0.9. Once S was determined, it was held fixed for all subsequent studies.

4.4.2 Validation Study

The effectiveness of the TCRS formulation in prospectively extracting turbidity free Raman information was tested on a separate set of 36 physical tissue models in which the concentration of the primary Raman scatterer was held constant. Similar to the calibration study, we found that the observed Raman spectra displayed appreciably different profiles (as shown in Fig. 4.4(a)), although in a clear medium, the uniform concentration of the Raman scatterer in each tissue phantom should give rise to the same Raman spectral profile. The mean spread index over all wavelengths for the observed Raman spectra was found to be 8.5%. On application of TCRS, the observed Raman spectra of the 36 phantoms tended to collapse onto one spectral profile as observed by the reduction of the mean spread index to 3.1% (Fig. 4.4(c)).

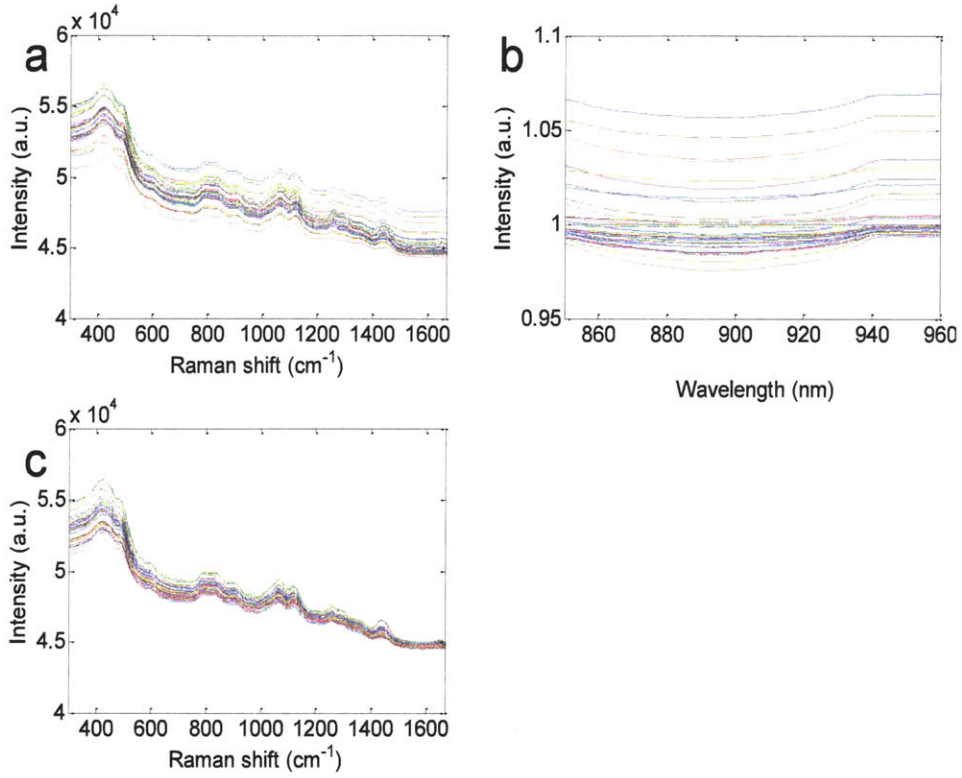


Fig. 4.4. (a) Observed Raman spectra and (b) normalized diffuse reflectance spectra of 20 representative tissue phantoms demonstrating typical spread. (c) Turbidity corrected Raman spectra showing that on the removal of turbidity-induced variations, Raman spectra of (a) tend to collapse onto a single spectral profile.

Thus, TCRS provides an average reduction of 64% over all wavelengths and up to 93% in the glucose fingerprint region (800-1400 cm^{-1}). It is notable that this substantial reduction is made in part possible by the fact that the proposed methodology is also able to correct for the turbidity-induced variations in the fluorescence in the tissue phantom, due to the fluorescent and diffusely reflected photons undergoing similar absorption and scattering events [14, 15].

Despite the significant improvement shown by the application of the TCRS methodology, the effect of absorption and scattering on Raman line shape and intensity is not completely removed.

There are two potential reasons behind the deviation from perfect coalescence observed in the spectral set here: (a) the presence of systematic noise (homoscedastic contributions such as detector noise and laser intensity fluctuations as well as heteroscedastic shot noise) and (b) effects of fluorescence quenching (photobleaching) of the sample holder (cuvette) over the duration of the experiments. The effects of photobleaching will be explored in greater detail in Chapter-6. Additionally, as the ratio of the scattering to the absorption coefficient decreases, the validity of the diffusion approximation becomes questionable - a topic of great interest in more absorbing media. We will also re-visit this issue, and the underlying theme of the diffusion approximation in absorbing media, in Chapter-5.

Further, OLS regression was used on the 36 phantom spectra to study the ability of TCRS to provide improved quantitative prediction capability. When the calibration model was used in conjunction with the uncorrected Raman spectra, the observed prediction error was 17.1%. In contrast, when the turbidity corrected Raman spectra were used, the prediction error dropped to 9.5%, a reduction of approximately 44%. To visualize this improvement, we show the boxplot of the ratio of predicted glucose concentrations (C_{obs}), obtained using uncorrected and turbidity corrected Raman spectra respectively, to the reference glucose concentrations (C_{ref}) in Fig. 4.5. The mean value of C_{obs}/C_{ref} is observed to be 0.87 and 0.98 for the uncorrected and turbidity corrected data respectively, thereby showing that the calibration model using turbidity corrected spectra is able to predict with a higher degree of accuracy. The standard deviation of the C_{obs}/C_{ref} values is found to be 0.12 and 0.09 for the uncorrected and turbidity corrected data, respectively, implying that the precision, i.e. the degree of reproducibility, of the concentration predictions is 25% higher for the turbidity corrected calibration model.

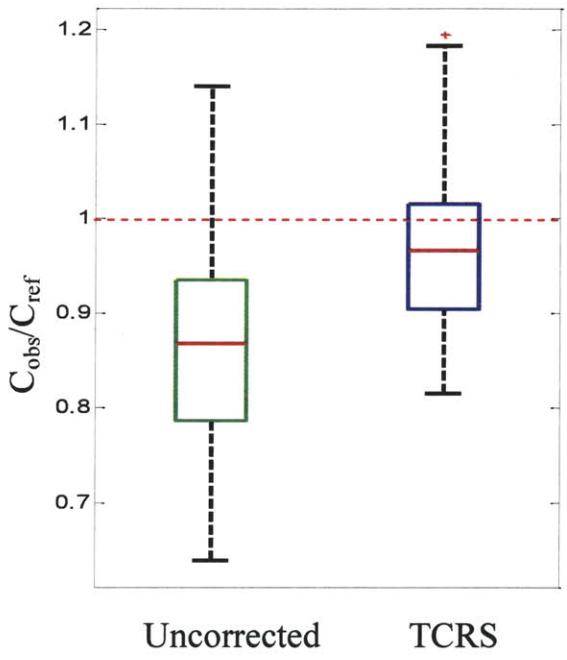


Fig. 4.5. Boxplot of ratio of predicted glucose concentrations (C_{obs}), to the reference glucose concentrations (C_{ref}) for uncorrected and turbidity corrected data. The dotted red line at 1 indicates the position where the observed glucose concentrations (extracted from OLS analysis) are equal to the reference glucose concentrations in the samples.

4.4.3 Prospective Prediction Study

The third study was used to compare and contrast the prospective prediction capability of the calibration models developed using uncorrected and turbidity corrected Raman spectra. In contrast to the calibration and intrinsic Raman studies described above, here the concentrations of glucose and creatinine were varied in a random manner over the 48 tissue phantoms. Creatinine was used as a spectral interferent to simulate an *in vivo* environment in which spectral interferents are present besides the analyte of interest (glucose). In order to simulate the conditions of a clinical study, where the spectra (or concentrations) of tissue constituents are not known *a priori*, an implicit calibration strategy was used to analyze the results of this study. The uncorrected and turbidity corrected Raman spectra were randomly divided into calibration and

prediction sets, containing 36 and 12 tissue phantom spectra, respectively. A standard PLS calibration procedure, into which the spectra and the corresponding glucose concentrations were input, was iterated 500 times to evaluate an average value of prediction error that is expected to be obtained in a general scenario. For both the calibration models, developed using uncorrected and turbidity corrected Raman spectra, respectively, 5 loading vectors were used in the PLS procedure. This was in accordance with the number of loading vectors that minimized the error in leave-one-out cross validation performed on the calibration set.

Figure 4.6 shows the boxplot of root-mean-square error of prediction (RMSEP) values obtained for glucose concentrations using both uncorrected and TCRS-corrected data. The mean prediction errors were 5.8 mM and 4.7 mM for the uncorrected and TCRS-corrected data, respectively, thereby demonstrating an approximately 20% reduction in prediction error on application of TCRS. This prospective prediction study thus illustrates that TCRS substantially removes concentration errors associated with turbidity. This demonstration represents an important step towards practical implementation of transcutaneous Raman scattering in human subjects.

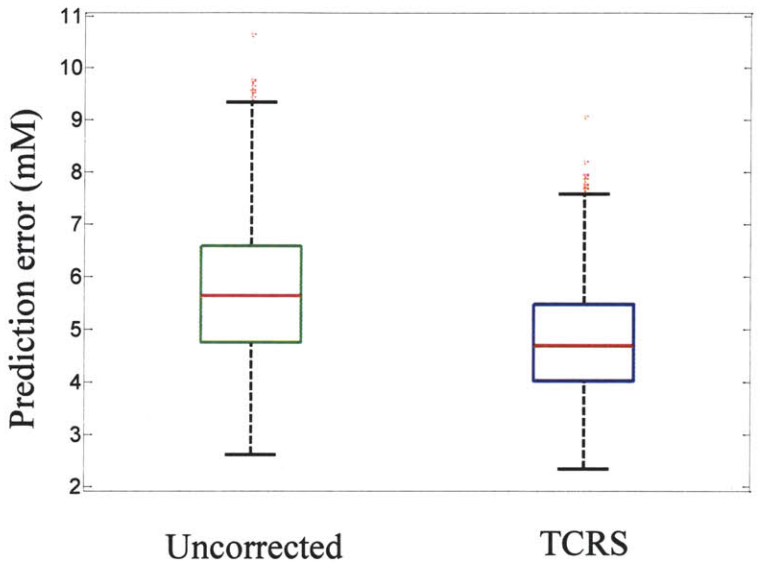


Fig. 4.6. Boxplot of RMSEP obtained for glucose concentrations from 500 iterations using uncorrected and TCRS-corrected data.

4.5 Summary

A central challenge in quantitative biological Raman spectroscopy, particularly applied to transcutaneous measurements, is overcoming the effects of subject-to-subject turbidity induced non-analyte specific variations. To account for these variations, we have developed a novel method, turbidity corrected Raman spectroscopy (TCRS), which is based on the photon migration approach and employs alternate acquisition of Raman and diffuse reflectance spectra. The new methodology developed here can be directly used in clinical studies, as it does not require the determination of tissue optical properties and the creation of tissue mimicking phantoms for calibration purposes. Our results clearly demonstrate that the TCRS methodology can extract intrinsic line shapes and intensity information from Raman spectra acquired in a turbid medium and provides a significant tool for prospective prediction studies. It is expected that TCRS will play a vital role in advancing Raman spectroscopy as a potential tool for

transcutaneous concentration measurements of blood analytes. The methodology presented here could be potentially extended to recalcitrant industrial process monitoring problems and analysis of contaminated fluid samples for environmental monitoring, where chemical quantification in turbid media is crucial.

Author's Post-Script

Since the implementation of this work and its publication in *Analytical Chemistry* [41], this approach has been cited for quantitative Raman spectroscopic correction in turbid media by several investigators, notably by Chaiken and co-workers [42, 43]. Specifically, more recent Monte Carlo model-based studies by the Helfmann group have validated our proposed correction methodology in diffusive regimes [44, 45].

Acknowledgements

Dr. Gajendra Pratap Singh was instrumental in modifying the experimental setup and performing the tissue phantom studies presented in this Chapter. We thank our colleagues Dr. Wei-Chuan Shih and Dr. Kate Bechtel for valuable discussions on this topic. Many of the figures and text were reproduced from the journal article [41] with permission from American Chemical Society (ACS).

Finally, from a personal standpoint, this chapter has a unique place in this thesis due to the in-depth involvement of Professor Michael Feld, from its beginning to its end. Indeed, while his creativity is inextricably linked to all of the work presented in this dissertation, he did not get to see the final outcomes of several of his ideas. Thus, it is with a mixture of emotions that we recall how he edited every last word of this specific work in his inimitable style.

4.6 References

(1) Berger, A. J.; Koo, T. W.; Itzkan, I.; Horowitz, G.; Feld, M. S. *Applied Optics* **1999**, 38, 2916-2926.

(2) Enejder, A. M. K.; Koo, T. W.; Oh, J.; Hunter, M.; Sasic, S.; Feld, M. S.; Horowitz, G. L. *Optics Letters* **2002**, 27, 2004-2006.

(3) Lambert, J. L.; Pelletier, C. C.; and Borchert, M. *Journal of Biomedical Optics* **2005**, 10, 1-8.

(4) Pelletier, C. C.; Lambert, J. L.; and Borchert, M. *Applied Spectroscopy* **2005**, 59, 1024-1031.

(5) Enejder, A. M. K.; Scecina, T. G.; Oh, J.; Hunter, M.; Shih, W.-C.; Sasic, S.; Horowitz, G.; Feld, M. S. *Journal of Biomedical Optics* **2005**, 10, 031114.

(6) Chaiken, J.; Finney, W.; Knudson, P. E.; Weinstock, R. S.; Khan, M.; Bussjager, R. J.; Hagrman, D.; Hagrman, P.; Zhao, Y. W.; Peterson, C. M.; Peterson, K. *Journal of Biomedical Optics* **2005**, 10, 031111.

(7) Chaiken, J.; Finney, W. F.; Yang, X.; Knudson, P. E.; Peterson, K. P.; Peterson, C. M.; Weinstock, R. S.; Hagrman, D. *Proceedings of SPIE* **2001**, 4254, 216-227.

(8) Khalil, O. S. *Clinical Chemistry* **1999**, 45, 165-177.

(9) Khalil, O. S. *Diabetes Technology & Therapeutics* **2004**, 6, 660-697.

(10) Arnold, M. A.; Small, G. W. *Analytical Chemistry* **2005**, 77, 5429-5439.

(11) Cote, G. L.; Lec, R. M.; Pishko, M. V. *IEEE Sensors Journal* **2003**, 3, 251-266.

(12) Keijzer, M.; Richards-Kortum, R. R.; Jacques, S. L.; Feld, M. S. *Applied Optics* **1989**, 28, 7.

(13) Durkin, A. J.; Jaikumar, S.; Ramanujam, N.; Richards-Kortum, R. *Applied Optics* **1994**, 33, 10.

(14) Wu, J.; Feld, M. S.; Rava, R. P. *Applied Optics* **1993**, 32, 3585-3595.

(15) Zhang, Q. G.; Muller, M. G.; Wu, J.; Feld, M. S. *Optics Letters* **2000**, 25, 1451-1453.

(16) Gardner, C. M.; Jacques, S. L.; Welch, A. J. *Applied Optics* **1996**, 35, 13.

(17) Patterson, M. S.; Pogue, B. W. *Applied Optics* **1994**, 33, 1963-1974.

(18) Zhadin, N. N.; Alfano, R. R. *Journal of Biomedical Optics* **1998**, 3, 16.

(19) Biswal, N. C.; Gupta, S.; Ghosh, N.; Pradhan, A. *Optics Express* **2003**, 11, 3320-3331.

(20) Cerussi, A. E.; Maier, J. S.; Fantini, S.; Franceschini, M. A.; Mantulin, W. W.; Gratton, E. *Applied Optics* **1997**, 36, 9.

(21) Finlay, J. C.; Foster, T. H. *Applied Optics* **2005**, 44, 17.

(22) Gupta, S.; Raja, V. L. N. S.; Pradhan, A. *Applied Optics* **2006**, 45, 9.

(23) Waters, D. N. *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy* **1994**, 50, 1833-1840.

(24) Nickell, S.; Hermann, M.; Essenpreis, M.; Farrell, T. J.; Kramer, U.; Patterson, M. S. *Physics in Medicine and Biology* **2000**, 45, 2873-2886.

(25) Tsuboi, M. *Journal of Biomedical Optics* **2002**, 7, 435-441.

(26) Aarnoutse, P. J.; Westerhuis, J. A. *Analytical Chemistry* **2005**, 77, 1228-1236.

(27) Kuba, S.; Knozinger, H. *Journal of Raman Spectroscopy* **2002**, 33, 325-332.

(28) Tinnemans, S. J.; Kox, M. H. F.; Nijhuis, T. A.; Visser, T.; Weckhuysen, B. M. *Physical Chemistry Chemical Physics* **2005**, 7, 211-216.

(29) Matousek, P.; Clark, I. P.; Draper, E. R.; Morris, M. D.; Goodship, A. E.; Everall, N.; Towrie, M.; Finney, W. F.; Parker, A. W. *Applied Spectroscopy* **2005**, 59, 393-400.

(30) Shih, W. C.; Bechtel, K. L.; Feld, M. S. *Optics Express* **2008**, 16, 12726-12736.

(31) Bechtel, K. L.; Shih, W. C.; Feld, M. S. *Optics Express* **2008**, 16, 12737-12745.

(32) Bonner, R. F.; Nossal, R.; Havlin, S.; Weiss, G. H. *Journal of the Optical Society of America* **1987**, 4, 10.

(33) Nossal, R.; Bonner, R. F.; Weiss, G. H. *Applied Optics* **1989**, 28, 7.

(34) Enejder, A. *Light scattering and absorption in tissue - models and measurements*, Ph.D. Thesis, Lund University, Sweden, **1997**.

(35) Wang, L.V.; Wu, H.-I., *Biomedical Optics: Principles and Imaging*, John Wiley & Sons, Hoboken, New Jersey, **2007**.

(36) Cheong, W.; Prahl, S.; Welch, A. J. *IEEE Journal of Quantum Electronics* **1990**, 26, 19.

(37) Tuchin, V. V.; Society of Photo-optical Instrumentation Engineers. *Tissue optics : light scattering methods and instruments for medical diagnosis*; SPIE Press: Bellingham, Wash., **2000**.

(38) Zonios, G.; Perelman, L. T.; Backman, V. M.; Manoharan, R.; Fitzmaurice, M.; Van Dam, J.; Feld, M. S. *Applied Optics* **1999**, 38, 6628-6637.

(39) Wold, S.; Martin, H.; Wold, H. *Lecture Notes in Mathematics,* Springer-Verlag: Heidelberg, **1983**.

(40) Flock, S. T.; Jacques, S. L.; Wilson, B. C.; Star, W. M.; Vangemert, M. J. C. *Lasers in Surgery and Medicine* **1992**, 12, 510-519.

(41) Barman, I; Singh, G. P.; Dasari, R. R.; Feld M. S.; *Analytical Chemistry*, **2009**, 81, 4233-4240.

(42) Chaiken, J; Goodisman, J; *Journal of Biomedical Optics*, **2010**, 15, 037007.

(43) Chaiken, J; Deng, B; Bussjager, R. J; Shaheen, G.; Rice, D.; Stehlik, D.; Fayos, J.; *Review of Scientific Instruments*, **2010**, 81, 034301.

(44) Reble, C.; Andree, S; Eichler, H.J.; Gersonde, I.; Helfmann, J; *Journal of Biomedical Optics*, **2010**, 15, 037016.

(45) Reble, C; Gersonde, I.; Lieber, C.A.l Helfmann, J.; *Biomedical Optics Express*, **2011**, 2(3), 520-532.

CHAPTER 5

IMPACT OF NON-LINEARITIES IN BIOLOGICAL RAMAN SPECTROSCOPY

Sample-to-sample variability as well the variability over time has proven to be a major challenge in achieving calibration maintenance and transfer in quantitative biological Raman spectroscopy. Multiple morphological and optical parameters, such as tissue absorption and scattering, physiological glucose dynamics and skin heterogeneity, vary significantly in a human population introducing non-analyte specific features into the calibration model. In this chapter, we show that fluctuations of such parameters in human subjects introduce curved (non-linear) effects in the relationship between the concentrations of the analyte of interest and the mixture Raman spectra.

To account for these curved effects, we propose the use of support vector machines (SVM) as a non-linear regression method over conventional linear regression techniques such as partial least squares (PLS). Using transcutaneous blood glucose detection as an example, we demonstrate that application of SVM enables a significant improvement in cross-validation prediction accuracy over PLS when measurements from multiple human volunteers are employed in the calibration set. Furthermore, we investigate the underlying cause of such curved effects and focus on sample turbidity principally due to its ability to distort intrinsic spectra, as discussed previously in Chapter-4. Using the same set of physical tissue models with randomized analyte concentrations and varying turbidities, we demonstrate that the fluctuations in turbidity alone causes curved effects, stemming from non-linear intensity scaling and peak width distortion (broadening/sharpening). As such, these distortions cannot be directly processed using

a linear model and have to be either explicitly corrected or must be modeled using non-linear regression techniques.

5.1 Significance

Tissue turbidity, as outlined in Chapter - 4, presents a formidable challenge in achieving successful calibration maintenance and transfer. In addition, other factors such as tissue autofluorescence and its quenching over time (Chapter - 6) and physiological lag between the blood and interstitial fluid glucose levels (Chapter - 7) also affect the robustness of the calibration model, i.e. the ability of the calibration model to predict on the same sample over time or/and on other samples not included in the original calibration dataset. Furthermore, any spurious correlation introduced into the calibration model will adversely affect its prospective prediction capability. Over the years, investigators have developed various correction schemes, incorporating changes in system design as well as spectral processing, to account for specific sources of variability (such as methods to correct for the turbidity-induced changes in sampling volume discussed earlier).

Nevertheless, given the broad range of such sources of variance, it is crucial that the calibration model is robust to non-analyte specific variations in the spectral data. Unfortunately, conventional implicit multivariate calibration (MVC) methods are prone to be misled by spurious effects such as system drift and covariance among constituents. To provide a more enhanced robustness than that afforded by the conventional calibration techniques [1], such as partial least squares (PLS) and principal component regression (PCR), researchers have previously developed hybrid MVC methods, which combine specific features of explicit and implicit calibration [2, 3], such as incorporation of prior information (e.g. pure spectrum of the analyte of interest) into the implicit calibration framework [4, 5].

Despite the improvements provided by the hybrid methods, all of the above schemes are limited by the underlying assumption that the relationship between the spectra and the property of interest (analyte concentration) is linear. In general, the total spontaneous Raman signal collected from a sample is the incoherent addition of the signals from individual molecules so that the acquired spectrum is linear in the concentration. This is also a reasonable first approximation for tissue Raman spectra as it has been previously shown that the Raman spectrum of a mixture sample can be treated as a linear superposition of the mixture's component spectra [6, 7]. However, the assumption of linearity *may* fail under the influence of fluctuations in process and system variables, such as changes in temperature and presence of sample turbidity. It is worth mentioning that weak non-linearities can be modeled by the conventional MVC methods by retaining larger number of factors than is necessitated by the chemical rank of the system – thereby also risking the inclusion of irrelevant sources of variance and noise in the calibration model. Stronger non-linear effects clearly present greater risks of creating spurious models that are incapable of prospective prediction.

Alternatively, the curved effects could be modeled by non-linear methods. Support vector machines (SVM) provide a relatively new class of such methods, which can handle ill-posed problems and lead to unique global models. Since its initial formulation by Vapnik and co-workers [8, 9], SVM has been used extensively for classification problems in bioinformatics [10] and chemometrics [11]. More recently, SVM has been extended to perform non-linear regression models capable of quantitative prediction. For example, SVM calibration models have exhibited excellent potential for near-infrared (NIR) absorption-based concentration prediction in chemical mixtures, even when the acquired spectra are affected by temperature fluctuations in a non-linear manner (i.e. peak distortions and peak frequency shifts) [12].

In this chapter, we first study the relationship between glucose concentrations and Raman spectral datasets; previously acquired from multiple human volunteers during the performance of oral glucose tolerance tests [13]. PLS implementation on the human subject dataset clearly demonstrates the presence of curved effects in the aforementioned relationship. To account for the deviation from linearity, we employ non-linear SVM calibration. We observe that SVM models provides a significant improvement over the PLS models with a reduction in prediction error by at least 30%, when multiple human subject datasets are included in the analysis. Nevertheless, making conclusions based on the human subject studies alone is perilous, given the uncontrolled nature of the changes in both process and physiological variables.

In the second part of this chapter, we hypothesize that one source of the observed non-linearity in the human subject dataset is the variations in tissue turbidity. To test our hypothesis in a controlled environment, we perform a set of physical tissue model (tissue phantom) studies with varying turbidities and randomized concentrations of the analyte of interest (glucose) and spectral interferents. Analysis of the tissue phantom data also reveals that SVM outperforms PLS, and PLS used in conjunction with turbidity corrected Raman spectroscopy, TCRS (presented in Chapter-4). This suggests the ability of SVM to provide an accurate calibration model, even in the presence of such non-analyte specific variances. Taken together, these results have important implications for clinical translation of Raman spectroscopy-based glucose monitoring and for other biological applications where sample-to-sample variability could lead to significant non-linear relationships between the property of interest and the measured spectra.

5.2 Theory of Support Vector Regression (SVR)

The theory of SVM and its extension to regression analysis has been described in detail elsewhere in the literature [14, 15]. For orientation, the primary equations of ε-SVM (epsilon-

SVM) regression are provided here. The central idea of this method is to find a function *f(x)* that has at most ε deviation from the actually obtained targets for all the calibration data (epsilon-tube criterion), where the calibration data is constituted by a set of regressors (x) and the corresponding dependent variables (y). (For the considered case of spectroscopic calibration, x and y denotes the spectra and concentration data, respectively.) In other words, a penalty is introduced for those data points which are far away from the predicted line (given by Eq. (5.1)) - but no such penalty is imposed if the point lies within some "epsilon-tube" of the predicted line.

For linear functions *f(x)* (i.e. for linear regression), this predicted line can be written as:

$$f(x) = \langle w, x \rangle + b \tag{5.1}$$

where w is the regression coefficient, b is the bias and <.,.> denotes the inner product of the vectors (w, $x \in R^d$ and $b \in R$, d is the dimension of the regressors, i.e. the number of spectroscopic data channels). In addition to the epsilon-tube criterion, the optimal function should be as "flat" as possible, i.e. where the magnitude of w is small. Large regression coefficients tend to reduce the generalization ability of the model due to the associated risk of overfitting. In support vector regression, a Lagrangian problem is formulated that balances this criterion with the minimization of the regression coefficients (since large regression coefficients are prone to overfitting). These two criteria can be mathematically combined into the following cost minimization problem for the calibration (training) data:

$$Q_{SVM} = \frac{1}{2}\|w\|^2 + C\sum_i (\xi_i + \xi_i^*) \tag{5.2}$$

subject to

$$y_i - \langle w, x_i \rangle - b \le \varepsilon + \xi_i \quad \forall i$$

$$\langle w, x_i \rangle + b - y_i \le \varepsilon + \xi_i^* \quad \forall i$$

$$\xi_i, \xi_i^* \qquad \geq 0 \qquad \forall i$$

where Q_{SVM} is the primal objective function that needs to be minimized, and ξ_i, ξ_i^* are slack variables introduced to allow for a "soft margin" loss function. The value of the constant C gives the trade-off between the first term, which minimizes the norm of w, and the second term, which takes into account the regression error for the training data. The slack variables are incorporated to allow the system to handle the otherwise infeasible error constraints of this quadratic optimization problem. Eq. (2) can now be re-written in the form of a Lagrangian function L_{SVM} as:

$$L_{SVM} = \frac{1}{2}\|w\|^2 + C\sum_i(\xi_i + \xi_i^*) - \sum_i(\eta_i\xi_i + \eta_i^*\xi_i^*) \tag{5.3}$$

$$-\sum_i \alpha_i(\varepsilon + \xi_i - y_i + \langle w, x_i\rangle + b) - \sum_i \alpha_i^*(\varepsilon + \xi_i^* + y_i - \langle w, x_i\rangle + b)$$

where η_i, η_i^*, α_i and α_i^* are the Lagrange multipliers for the inequality constraints in Eq. (5.2). From Eq. (5.3), it is evident that each training data point has a separate set of Lagrange multipliers, which decide the impact of the point on the final solution. Specifically, if a data point lies inside the epsilon-tube, it has no effect on the final solution. This results in an overall sparse solution where the final model depends only on a fraction of the training data. The partial first derivatives of this Lagrangian function (L_{SVM}) are then set to zero to obtain the final SVM model. In particular, equating the partial derivative of the Lagrangian function with respect to w to zero yields the support vector expansion:

$$w = \sum_i(\alpha_i - \alpha_i^*)x_i \tag{5.4}$$

Combining Eq. (5.1) & (5.4), we get:

$$f(x) = \sum_i (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b \tag{5.5}$$

An important advantage of this approach is that it can be readily extended for handling non-linear regression. It has been shown that the inner product $<x_i,x>$ can be replaced by a kernel function $K(x_i,x)$ as long as the kernel function satisfies Mercer's conditions [15]. The resultant regression model is given by Eq. (5.6):

$$f(x) = \sum_i (\alpha_i - \alpha_i^*) K(x_i, x) + b \tag{5.6}$$

The non-linear transformation is performed via the kernel function $K(x_i,x)$. The most common choices of non-linear kernels are the radial basis function ($\exp(-\|x_i-x_j\|^2/2\sigma^2)$, where σ^2 is the width of the RBF) and the polynomial function ($<x_i,x_j>^d$, where d is the polynomial degree). The kernel parameters, ε and C are optimized by the user as per the training dataset. Further details of our SVM implementation procedure are discussed in Chapter-5.4 below.

Finally, it is worth mentioning that, in addition to the sparseness advantage (Arising from the epsilon-tube criterion), the SVM formulation also provides unique optimal solutions in contrast to other non-linear regression techniques such as artificial neural networks (ANN), which suffer from multiple local minima [12]. Moreover, the reason that SVMs often outperform ANNs in practice is that unlike ANNs, SVMs are less prone to overfitting.

5.3 Experimental Methods

Experimental studies are undertaken to accomplish the following objectives: (1) compare the effectiveness of SVM and conventional PLS calibration models when multiple human subject datasets are incorporated in the analysis and (2) investigate the effect of turbidity-induced

sampling volume variations on the relationship between the reference concentrations and the acquired Raman spectra. To fulfill objective (1), multivariate analysis is undertaken on datasets obtained from our laboratory's previous clinical studies on human volunteers. In addition, the study is used to understand the nature of non-linearity introduced, if any, as multiple human volunteers are included in the calibration model. To accomplish objective (2), a tissue phantom study is designed where the concentrations of the analyte of interest and spectral interferents as well as the turbidities of the phantoms are randomly varied. The tissue phantom study provides a degree of control and specificity, which is difficult to implement in a human subject or animal model study. In the latter cases, fluctuations in other parameters including tissue fluorescence, skin heterogeneity and physiological glucose kinetics would obscure the effect of turbidity variation on the relationship between the spectra and concentration datasets. In this study, Raman and diffuse reflectance spectra are collected from the tissue phantoms for comparison of the PLS, SVM and turbidity corrected PLS calibration models. Correction for turbidity variation is performed in accordance with the TCRS procedure.

5.3.1 Human volunteer study

The acquisition of the clinical data from the human volunteers was originally described in one of our laboratory's previous publications [13]. Briefly, Raman spectra were acquired from the forearms of healthy Caucasian and Asian volunteers undergoing oral glucose tolerance tests (OGTT). An 830 nm diode laser was used as the Raman excitation source and a gold-coated half-paraboloidal mirror was employed to enhance the collection efficiency of the system. The collected light was fiber-coupled to a f/1.8 spectrograph for dispersion before spectral acquisition using a liquid nitrogen cooled deep depletion CCD. In accordance with a typical OGTT protocol, each volunteer was given a glucose-rich drink. Following this, spectra were collected every 5

minutes, with an acquisition time of 3 minutes, over a 2-3 hour measurement period. In addition, finger prick measurements of blood glucose levels were performed at 10 minute intervals using a HemoCue glucose analyzer.

5.3.2 Tissue phantom study

For this study, the experimental setup and the protocol of physical tissue model fabrication was identical to that described in Chapter-4. A set of 50 tissue phantoms were fabricated by pipetting out randomized concentrations of glucose (analyte of interest), creatinine (Raman spectral interferent), intralipid (anisotropic elastic scatterer) and India ink (absorber) into distilled water. The values of glucose and creatinine concentrations as well as the scattering and absorption coefficients were varied over the same ranges as described earlier. The ambient temperature in this study was carefully regulated to 73° ± 0.5°F to minimize the effects of temperature distortions.

5.4 Data Analysis

For both human volunteer and tissue phantom studies, conventional linear regression (PLS) and non-linear SVM regression methods are employed.

5.4.1 Human volunteer study

Of the 17 volunteers on whom the OGTT were successfully performed, 4 volunteers exhibited impaired glucose tolerance profiles. These 4 volunteer datasets are excluded from further analysis, given the risk of incorporating temporal spurious correlations between the nearly monotonic glucose profiles and spectral altering parameters (such as fluorescence quenching and system drift) [16]. Spectra from 355-1545 cm^{-1} are used in all data analysis after image curvature correction, vertical binning and cosmic ray removal. The datasets from the human volunteers are

directly input for development of the MVC calibration models without any data manipulation such as removal of fluorescence (which could potentially incorporate spectral artifacts, as discussed in Chapter-6).

The 13 volunteer datasets constituting a total of 350 data points (spectra and concentrations) are first analyzed using a PLS leave-one-out cross-validation (LOOCV) routine to construct a "global" calibration model. The purpose of the LOOCV process is to optimize the parameters, namely loading vectors for PLS and C and σ^2 for SVM. As in a typical LOOCV procedure, one data point is left out at a time from the calibration data and the developed model is used to compute the concentration of the left out data point. This procedure is then repeated until all data points have been left out in turn. The resultant root mean square error of cross validation (RMSECV) is used as a metric for tuning the parameters of the MVC models. For the PLS regression analysis, the number of loading vectors which provide the minimum RMSECV is first determined. Next, to reduce chances of overfitting, the calibration model corresponding to the minimum number of loading vectors, which gives a less than 5% deviation from the previously obtained minimum RMSECV, is selected. (Moreover, for all PLS analysis we satisfied the criterion that the number of calibration samples is at least 3 times larger than the rank of the calibration model, i.e. the number of loading vectors [17].) In addition to global calibration, PLS "local" calibration is performed by using the LOOCV protocol on each of the 13 volunteer datasets individually.

The calculations pertaining to the multivariate non-linear calibration are performed using a SVM MATLAB toolbox [18]. The RBF kernel is used for non-linear regression. Prior to SVM leave-one-out cross-validation, the Raman spectra are linearly scaled such that the intensity values were distributed between 0 and 1. The scaling is performed by dividing each spectrum by

the maximum intensity value across the entire spectral range under consideration. The scaling step is performed to avoid the dominance of specific pixel intensity values that reside in greater numeric ranges over those having smaller numeric values. The optimal model parameters C and σ^2 (i.e. the parameters that provide the smallest RMSECV) are obtained using a grid search algorithm in the range of 1 to 10000 (C) and 0.01-10 (σ^2), respectively. Based on the results of prior studies, the ε parameter is kept constant at 0.001. SVM regression is used to generate only a global model – in contrast to PLS implementation where a global as well as multiple local models are generated. In addition, a sample out cross validation study for SVM is performed by omitting each human subject's dataset in turn from calibration and using it for prospective prediction. For this study, we allow a slope and bias term to be fit for each subject (due to the limited number of samples available for calibration). The objective of this specific study is to explore the possibility of prospective prediction in human population.

5.4.2 Tissue phantom study

For the tissue phantom study, the experimental dataset, consisting of 50 samples in all, is randomly split into 36 samples for calibration set and 14 for prediction. PLS and SVM models are first generated on the calibration set based on a leave-one-out cross-validation procedure (similar to that mentioned above for the human subject study). For the SVM model, a grid search is performed over the following parameter space: 1 to 20000 (C) and 0.01-10 (σ^2); whereas for the PLS model the loading vectors are varied from 3 to 20. The developed calibration models are then used prospectively on the prediction set to determine the glucose concentrations. Overfitting in this study is avoided by invoking an independent prediction set (i.e. not present in the calibration model) to report the prospective errors (root-mean-square error of prediction, RMSEP). The creation of an independent prediction set is a standard approach to mitigating

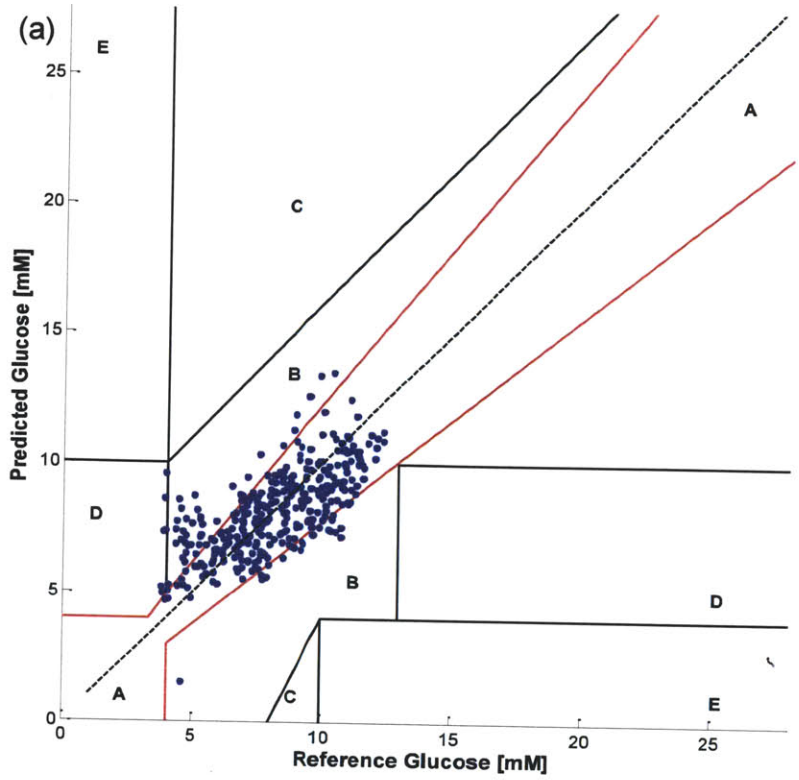
and/or testing for the presence of spurious correlations [1]. If the data of the calibration set is overfit with a large number of factors, its ability to predict prospective samples will be seriously compromised due to the incorporation of noise and spurious elements. The entire analytical procedure (i.e. splitting of the experimental dataset, followed by calibration and prediction) is iterated 100 times to evaluate an average prediction error value.

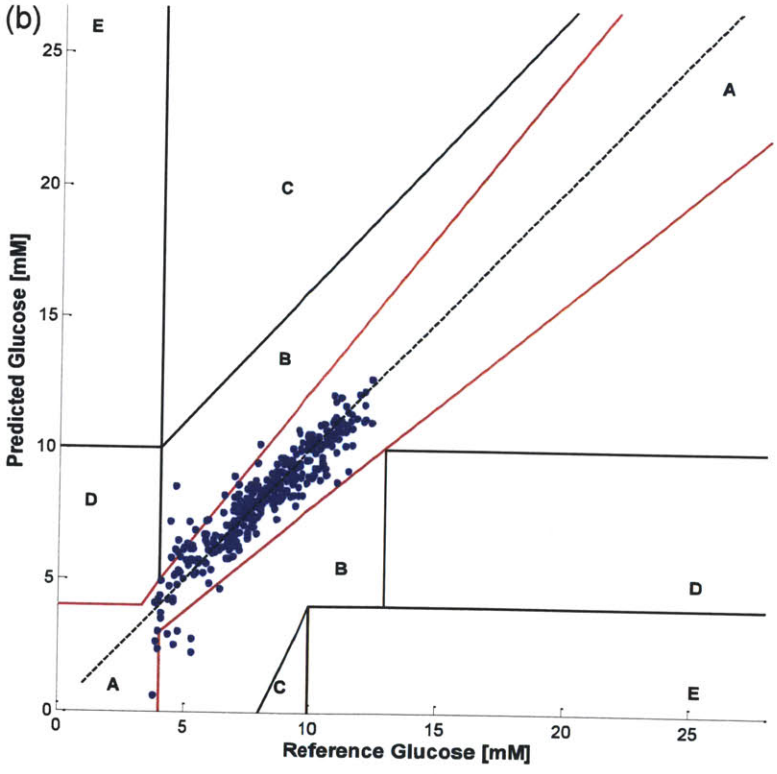
Additionally, as mentioned above, turbidity correction is undertaken on the same calibration and prediction datasets using the TCRS formalism. The Raman intensity at emission wavelengths is divided by the square root of the product of the diffuse reflectance intensities at the laser excitation and Raman emission wavelengths to obtain the "intrinsic" (turbidity-free) Raman spectra. The resulting intrinsic Raman spectra are then analyzed using the PLS procedure. It is worth mentioning that in each of the 100 iterations, the calibration and prediction sets for PLS (raw and TCRS-applied) and SVM are kept identical.

5.5 Results and Discussion

5.5.1 Human volunteer study

A representative set of tissue Raman spectra and the corresponding blood glucose concentration profile acquired from one of the human volunteers undergoing an OGTT can be seen in Chapter-3. Glucose cross-validation results on the human volunteer datasets - using PLS global (a), PLS local (b), and SVM global (c) - are shown in Fig. 5.1 plotted on the Clarke error grid [19], a widely used method for evaluating the clinical usefulness of glucose predictions. (Predictions in zones A and B are considered acceptable, and predictions in zones C, D, and E are potentially dangerous if used in clinical judgment.)


(a)
E
A
C
B
D
B
D
A
C
E
Predicted Glucose [mM]
Reference Glucose [mM]
0
5
10
15
20
25


(b)
E
A
C
B
D
B
D
A
C
E
Predicted Glucose [mM]
Reference Glucose [mM]
0
5
10
15
20
25

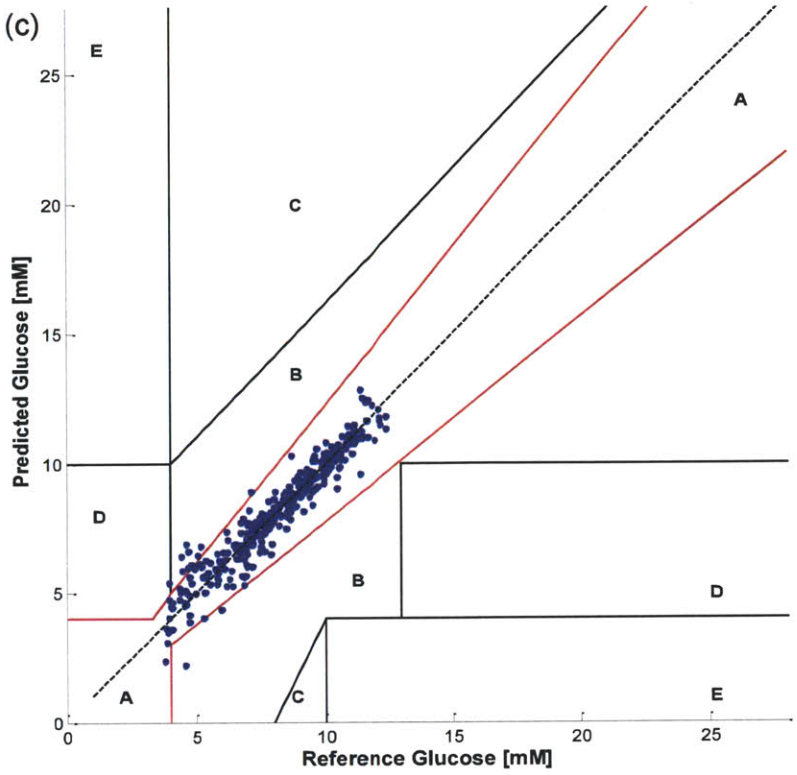


Fig. 5.1 Blood glucose predictions of different approaches shown on the Clarke Error Grid: (a) PLS global; (b) PLS local; and (c) SVM global.

From the figure it is evident that application of the PLS global model exhibits significantly larger errors and worse correlations (RMSECV = 28 mg/dl, R^2 = 0.67) compared to the SVM global model (RMSECV = 10.87 mg/dl, R^2 = 0.96). In fact, the glucose predictions generated with the SVM global model show a better match to the reference glucose concentrations, even compared to those generated using PLS local calibration (RMSECV = 15.45 mg/dl, R^2 = 0.92). In other words, the SVM global calibration provides an improvement of *ca.* 61% and 30% over PLS global and PLS local calibration, respectively. It is also worth mentioning that the beneficial impact of nonlinear regression is particularly pronounced in the low analyte concentration region (when the reference glucose concentration is in the range of 4-6 mM). This represents an improvement in the prediction accuracy, or reduction in bias - rather than an improvement of

precision, which is a direct function of systematic noise. The reduction in bias can potentially be attributed to the robust nature of SVM models due to the regularization process, with regard to "errors" in reference spectra (i.e. spectral distortions) or/and reference concentrations, as explained below.

To understand the root cause of this disparity in model performances, the development of the PLS global model is investigated. We find that the PLS model employs increasingly larger number of more loading vectors as more datasets are incorporated in the analysis (Fig. 5.2).

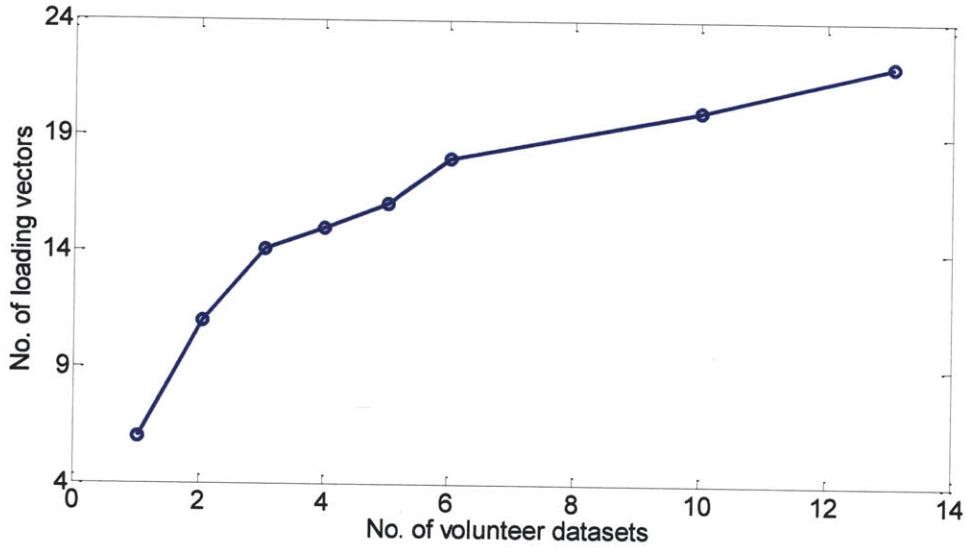


Fig. 5.2 Plot of the optimal number of loading vectors (as determined by the minimum RMSECV) as a function of the number of volunteer datasets included in the PLS analysis.

Clearly, this rise in the number of optimal loading vectors (with the inclusion of additional volunteers) cannot be attributed to the chemical rank of the system as human skin tissue has a unique and consistent set of principal chemical components [13]. While a small increase in the number of loading vectors over that obtained for a single individual may be explained by the presence of instrumental drift and variable baselines, an increase from 6 LV for 1 human subject

to 22 LV for 13 human subjects cannot be ascribed to these factors. This behavior implies the presence of significant curved effects in the relationship between Raman spectra and the reference glucose concentrations. Obviously, the retention of more regression factors in the PLS global model than that warranted by the chemical rank comes at the cost of including irrelevant sources of variance and noise (as verified by Fig. 5.3). In contrast to PLS global calibration, the PLS local calibration models developed individually on each volunteer dataset only employ 5-8 loading vectors. Nevertheless, the PLS local models themselves are susceptible to spurious correlations in the individual dataset that prevents adequate calibration maintenance (on the same subject over time). For the SVM global model, the optimal combination of C and σ^2 was established to be 1000 and 0.05, respectively. Importantly, this specific combination of C and σ^2 resulted in a less than 2% change in the calculated RMSECV, regardless of the number of volunteers included in the analysis, showing the relative constancy of the optimal parameter combination in clear contrast to the PLS global model.

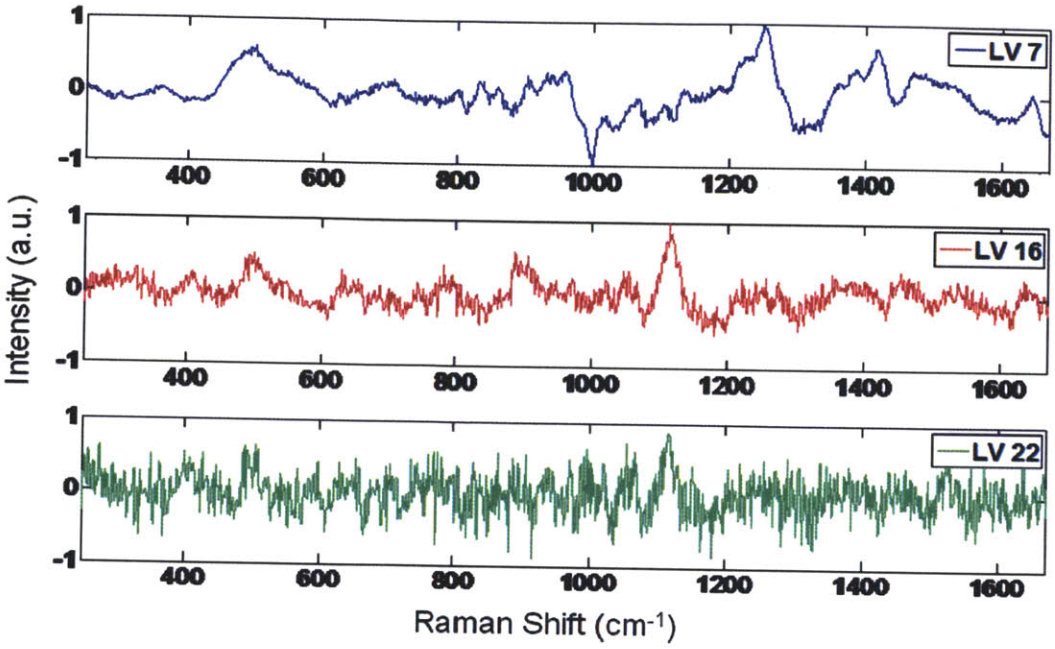


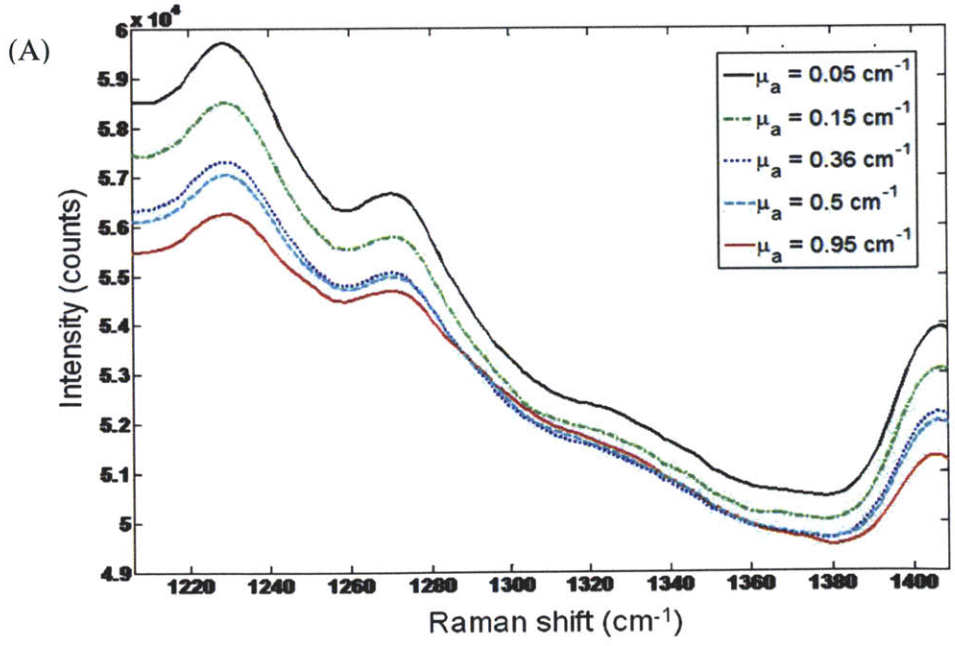
Fig. 5.3 Comparison of the regression vectors for the PLS global model corresponding to 7, 16 and 22 regression factors (loading vectors), respectively. While 22 loading vectors provide the least error in cross-validation for 13 volunteers (Fig. 5.2), the visibly noisy nature of the corresponding b-vector is unlikely to provide accurate concentration estimates on data that has not been part of the original calibration.

While kernel-based support vector regression enables the non-linear modeling between the spectra and concentration datasets, the sources of the non-linear interferences in this relationship still need to be carefully examined. Some of these non-linear interferences can be attributed to fluctuations in process and sample variables, which are not specific to the analyte of interest but alter the spectral measurements. The non-linear influence of external process variables such as temperature on vibrational spectra is well-known [20, 21] and is particularly relevant when measurements are performed in an actual process environment where well-controlled laboratory conditions are not achievable.

It is also important to note that one of the primary challenges in demonstrating calibration transfer is the establishment of causation between the glucose levels and the optical signals (i.e. whether the source of the signal is truly a change in glucose). For all calibration models (linear or non-linear), spurious effects may arise from different sources that correlate with changes in glucose, especially in human subject studies. Some of these sources may stem from the glucose tolerance test itself, e.g. due to a small physiological change or a change in room temperature which could correlate with observed glucose changes. In order to more rigorously understand the underlying potential causes of such curved effects in the spectra-concentration relationship, we employ a tissue phantom study, where such non-analyte specific sources can be eliminated due to the controlled environmental settings.

5.5.2 Tissue phantom study

Effect of turbidity on Raman spectral distortions



(B)

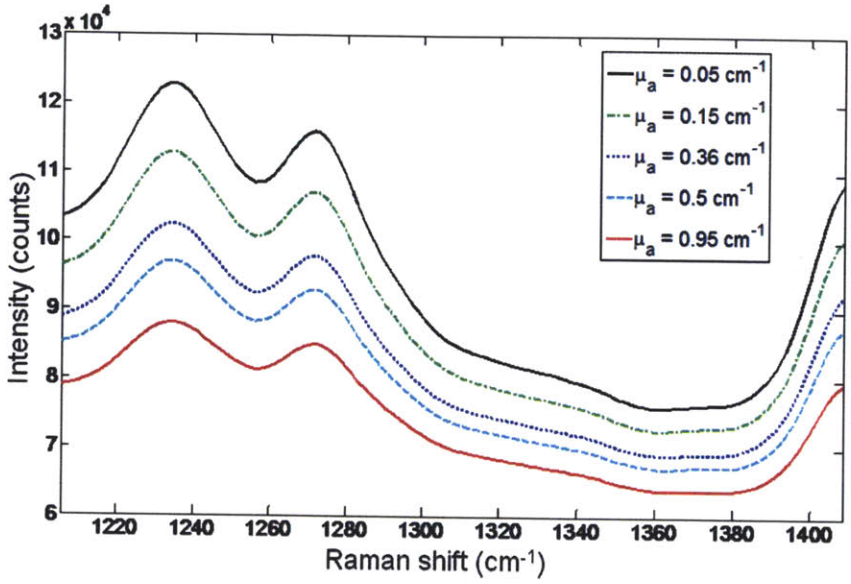


Fig. 5.4 Observed Raman spectra from two sets of five tissue phantoms showing the typical spread. Each set has same scattering coefficient but different absorption coefficients. The scattering coefficient (μ_s) of the tissue phantoms in (A) (top panel) and (B) (bottom panel) are 24 and 130 cm^{-1}, respectively. The absorption coefficients are marked in the legend.

Figure 5.4 represents spectra acquired from tissue phantoms having a fixed concentration of the analyte of interest (Raman scatterer) but different (elastic) scattering and absorption coefficients. Panel (A) and (B) shows the spectra obtained at two different values of the scattering coefficient of the tissue phantom, μ_s = 24 cm^{-1} and μ_s = 130 cm^{-1}, respectively. These spectra were acquired as part of a 49 tissue phantom study where the Raman scatterer concentration was kept constant and the phantom absorption and scattering coefficients were varied over the range of values observed in human skin tissue. First, it is clear that the presence of differing amounts of turbidity causes a significant spread in the observed spectral profiles in both (A) and (B), as seen previously in Chapter-4. Second, we observe that the observed spectra from the tissue phantoms having the higher scattering coefficient (B) show an approximately linear scaling over the entire

wavelength range mediated by the value of the absorption coefficient in the tissue phantom. In contrast, at the smaller scattering value of the tissue phantoms in (A), the observed Raman spectra exhibit a more complex scaling behavior - which is more pronounced at the higher absorption values. Indeed at higher absorption (μ_a = 0.36, 0.5 and 0.95 cm^{-1}) to scattering ratios, we observe substantive spectral overlap which means there is a peak distortion effect, in addition to the scaling.

These findings can be attributed to the relative interplay of the elastic scattering and absorption in the tissue phantoms. It has been previously observed that when the reduced scattering coefficient ($\mu_s' = \mu_s*(1-g)$) is significantly larger than the absorption coefficient (diffusive regime), the observed spectrum (fluorescence or Raman) could be linearly scaled by the diffuse reflectance intensities to extract the intrinsic spectrum of the sample [22, 23]. However, this simple relationship fails when the diffusion approximation is no longer valid, as pointed out by Zhang *et al.* [24]. Our observations in Fig. 5.4 are in line with such an understanding.

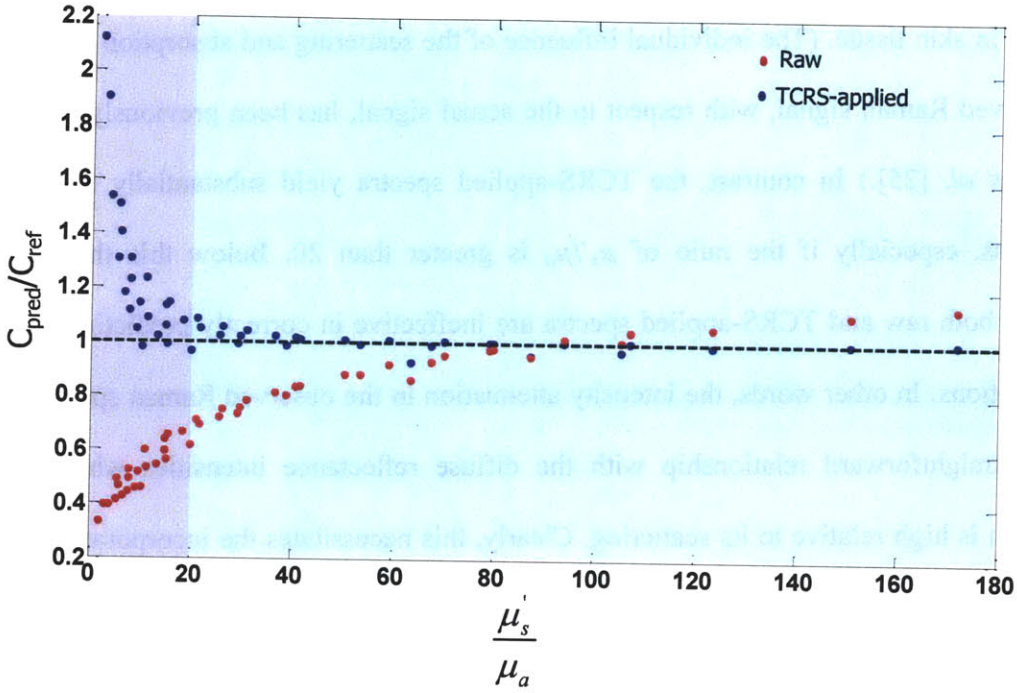


Fig. 5.5 Ratio of predicted to actual values of the analyte of interest at constant concentration plotted as a function of μ_s'/μ_a. The red circles give the OLS prediction values using acquired spectra. The blue circles indicate the prediction values obtained with OLS on TCRS-applied spectra.

This fact is further borne out when ordinary least squares (OLS) is performed on the raw and TCRS-applied spectral dataset using the model components as the basis spectra. Figure 5.5 plots the ratio of the observed concentrations to the reference concentrations of the Raman scatterer (C_{obs}/C_{ref}) in the 49 tissue phantoms as a function of μ_s'/μ_a in the tissue phantoms. The dotted black line at $C_{obs}/C_{ref} = 1$ indicates the position where the glucose concentrations computed from OLS analysis are equal to the reference glucose concentrations in the samples. The red circles indicate the values obtained with raw spectral analysis and the blue circles correspond to those computed with TCRS-applied spectra. It is clear that the OLS analysis of the raw spectra provides incorrect estimates of the concentrations over the whole range of turbidity values

observed in skin tissue. (The individual influence of the scattering and absorption parameters on the observed Raman signal, with respect to the actual signal, has been previously discussed by Bechtel *et al.* [25].) In contrast, the TCRS-applied spectra yield substantially more accurate predictions, especially if the ratio of μ_s'/μ_a is greater than 20. Below this threshold ratio, however, both raw and TCRS-applied spectra are ineffective in correctly predicting the analyte concentrations. In other words, the intensity attenuation in the observed Raman spectra does not have a straightforward relationship with the diffuse reflectance intensities when the tissue absorption is high relative to its scattering. Clearly, this necessitates the incorporation of a non-linear methodology to account for the high absorption regime turbidity variations.

Our observations here are also analogous to the findings from Wülfert's seminal study on the effect of temperature on NIR absorption spectra [20]. In this study, the researchers observed that variations in temperature translate via the changes in intermolecular forces into non-linear modifications in the vibrational spectra. Using numerical simulations and experiments with alcohol mixtures, it was shown that this effect can be manifested in terms of a change in peak area, a frequency shift or changing width of the peaks. For our turbidity study, the change in peak area is the predominant concern followed by the changing width of the peaks. It is worth mentioning that the changing peak width is of greater importance in studying fluorescence, which exhibits broader features, than Raman, which has significantly sharper peaks and therefore is less affected by the small local variations in tissue absorption especially in the NIR region. Put another way, the observed Raman spectra primarily undergo scaling distortions and minor local distortions as opposed to fluorescence spectra, which may undergo significant shape variations as well [26].

Consequences of turbidity-induced variations on multivariate calibration models

From Fig. 5.5, we get a clear indication about the detrimental impact of turbidity variation on the concentration estimate of the analyte of interest. Intuitively, one can understand that if the absorption is high then less Raman photons are back-scattered leading to an underestimate on the true concentration of Raman scatterers. While TCRS corrects this by using diffuse reflectance measurements, we find that it is limited by the boundaries of the underlying diffusion approximation. Thus, one needs to incorporate some form of non-linear modeling when the tissue absorption is high relative to its scattering value. There are two approaches which can be used to this end: (i) introduction of a non-linear function which accurately encapsulates the changes in Raman spectra with respect to the diffuse reflectance spectra in the high absorption regime and (ii) introduction of a non-linear multivariate calibration model robust to sample turbidity-induced variations.

In the first approach (i), one can express the intrinsic Raman spectrum as a ratio of the observed spectrum to some pre-calibrated (non-linear) function of the diffuse reflectance spectrum (such as a power law or exponential function) [27, 28]. This function would, in turn, depend on the sample turbidity parameters and excitation-collection geometry. Unfortunately, this would necessitate the explicit determination of the optical properties and the performance of additional tissue phantom studies (where the reproducibility of the function *in vivo* would depend on the exactness of the fabricated tissue phantoms). Here, we propose an alternate approach (ii) by introducing non-linear kernel-based support vector regression, which may potentially be robust to the turbidity-induced variations. This approach can be developed on the human subject dataset itself without necessitating additional tissue phantom model studies or incorporating a broadband source.

For the tissue phantom study presented here, standard PLS procedure was carried out on the raw and TCRS-applied Raman spectra using an optimal of 5 loading vectors. Support vector regression was carried out only on the raw spectra with an optimal (C, σ^2) parameter combination of (10000, 1).

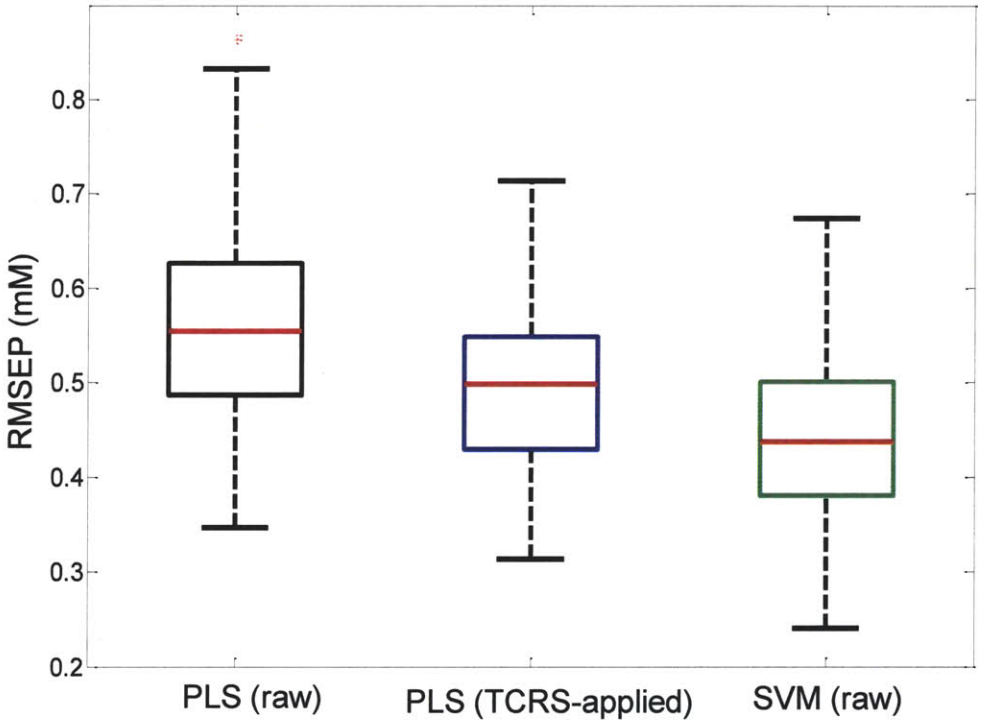


Fig. 5.6 Boxplot of RMSEP obtained for glucose concentrations from 100 iterations using PLS on acquired spectra (raw), PLS on TCRS-applied spectra and SVM on raw spectra.

Figure 5.6 shows the boxplot results of the root-mean-square error of prediction (RMSEP) values obtained for the glucose concentrations. The mean prediction errors were 0.55 mM, 0.5 mM and 0.44 mM for the PLS (raw), PLS (TCRS-corrected) and SVM (raw) data, respectively. This demonstrates a 20% reduction in prediction error on application of SVM in the raw case. Importantly, the SVM case improves the prediction accuracy even when PLS is employed on the TCRS-applied spectra by about 12%. This improvement in predictive ability of glucose was

observed largely for those tissue phantoms with high absorption values. Given that this study was performed under well-controlled laboratory conditions with temperature stabilization, we can conclude that SVM is able to deal with the curved effects introduced by the turbidity variations of the various tissue phantoms. Further, when SVM was applied to TCRS-corrected spectra for the tissue phantom study, we observed negligible alteration (<1%) from the results of the SVM application on the unprocessed spectra. This negligible change in resulting prediction errors can be attributed to the fact that SVM already accounts for the turbidity-induced non-analyte specific variances.

The enhanced robustness provided by SVM models, as compared to PLS, can be attributed to two primary factors in addition to its capability to model non-linear datasets. First, the assignment of relative weights (importance) of training samples using Lagrange multipliers in conjunction with the presence of the ε-insensitive loss function forces a strict distinction between important training samples and irrelevant ones. This provides the necessary flexibility to deal with outliers, which unduly influence conventional least squares methods [29, 30]. Second, the regularization term introduced in the formulation penalizes the large regression weights thereby ensuring that small variations in input concentrations and spectra do not lead to large variations in the regression model. Such an approach has been extensively used in ill-conditioned problems (e.g. Tikhonov regularization) where the noisy nature of the data necessitates a robust solution methodology [31].

Finally, it should be noted that although application of SVM enhances prediction accuracy and robustness, it remains a mathematical ("black-box") approach which does not specifically address the root cause of any particular curved effect such as those introduced by sample turbidity. Given the large number of variable parameters in non-invasive glucose studies in a

human population, we recommend that SVM should be used in conjunction with other physical approaches that specifically accounts for such variations.

5.6 Summary

Robustness of the spectroscopy-based calibration model to fluctuations of system and sample parameters is a critical first step towards successful prospective application. Linear MVC methods can often be misled by spurious correlations and covariance among constituents and tend to incorporate larger number of factors, especially in the presence of non-linearities. Additionally, large factor models are problematic from a SNR, calibration maintenance and calibration transfer perspective. In this chapter, we have proposed the application of non-linear support vector machine regression to enhance the robustness and prediction accuracy of the calibration model. We have demonstrated that SVM provides nearly 30% enhancement in prediction accuracy over PLS calibration models, when measurements from multiple human volunteers are considered. Furthermore, we have shown that variations in turbidity alone may introduce non-linearity into the spectra-concentration relationship, especially when tissue absorption is comparable to its scattering.

Acknowledgements

Dr. Narahara Chari Dingari's help in polishing the text and the figures is gratefully acknowledged. Chae-Ryon Kong verified the computations presented here and is currently extending the application of SVM to analysis of the data sets collected from our new human volunteer studies using the clinical Raman instrument. Many of the figures and text were reproduced from the journal article [32] with permission from American Chemical Society (ACS).

5.7 References

(1) Brereton, R. G. *Applied Chemometrics for Scientists*, John Wiley & Sons Ltd., Chichester, West Sussex, England; **2007**.

(2) Wentzell, P. D.; Andrews, D. T.; Kowalski, B. R. *Analytical Chemistry* **1997**, 69, 2299-2311.

(3) Haaland, D. M.; Melgaard, D. K. *Applied Spectroscopy* **2001**, 55, 1-8.

(4) Berger, A. J.; Koo, T. W.; Itzkan, I.; Feld, M. S. *Analytical Chemistry* **1998**, 70, 623-627.

(5) Shih, W. -C.; Bechtel, K.; Feld, M. S. *Analytical Chemistry* **2007**, 79, 234-239.

(6) Hanlon, E. B.; Manoharan, R.; Koo, T. W.; Shafer, K. E.; Motz, J. T.; Fitzmaurice, M.; Kramer, J. R.; Itzkan, I.; Dasari, R. R.; Feld, M. S. *Physics in Medicine and Biology* **2000**, 45, R1-R59.

(7) Manoharan, R.; Shafer, K.; Perelman, L.; Wu, J.; Chen, K.; Deinum, G.; Fitzmaurice, M.; Myles, J.; Crowe, J.; Dasari, R. R.; Feld, M. S. *Photochemistry and Photobiology* **1998**, 67, 15-22.

(8) Cortes, C.; Vapnik, V. *Machine Learning* **1995**, 20, 273-297.

(9) Vapnik, V. *The Nature of Statistical Learning Theory*, Springer-Verlag; **1995**.

(10) Ramaswamy, S.; Tamayo, P.; Rifkin, R.; Mukherjee, S.; Yeang, C.; Angelo, M.; Ladd, C.; Reich, M.; Latulippe, E.; Mesirov, J. P.; Poggio, T.; Gerald, W.; Loda, M.; Lander, E. S.; Golub, T. R. *Proc. Natl. Acad. Sci.* **2001**, 98, 15149-15154.

(11) Widjaja, E.; Lim, G. H.; An, A. *Analyst* **2008**, 133, 493-498.

(12) Thissen, U.; Ustun, B.; Melssen, W. J.; Buydens, L. M. C. *Analytical Chemistry* **2004**, 76, 3099-3105.

(13) Enejder, A. M. K.; Scecina, T. G.; Oh, J.; Hunter, M.; Shih, W.; Sasic, S.; Horowitz, G. L.; Feld, M. S. *Journal of Biomedical Optics* **2005**, 10, 031114-9.

(14) Scholkopf, B.; Smola, A. J. *Learning with Kernels*, MIT Press: Cambridge, MA; **2002**.

(15) Christianini, N.; Shawe-Taylor, J. *An introduction to support vector machines*, Cambridge University Press: New York; **2000**.

(16) Liu, R.; Chen, W.; Gu, X.; Wang, R. K.; Xu, K. *Journal of Physics D: Applied Physics* **2005**, 38, 2675-2681

(17) Qi, D.; Berger, A. J. *Applied Optics* **2007**, 46, 1726-1734.

(18) Chang, C. C.; Lin, C. -J., *LIBSVM: a library for support vector machines* **2001**.

(19) Clarke, W. L.; Cox, D.; Gonder-Frederick, L. A.; Carter, W.; Pohl, S. L. *Diabetes Care* **1987**, 10, 622-628.

(20) Wülfert, F.; Kok, W. T.; Smilde, A. K. *Analytical Chemistry* **1998**, 70, 1761-1767.

(21) Despagne, F.; Massart, D. L.; Chabot, P. *Analytical Chemistry* **2000**, 72, 1657-1665.

(22) Barman, I.; Singh, G. P.; Dasari, R. R.; Feld, M. S. *Analytical Chemistry* **2009**, 81, 4233-4240.

(23) Wu, J.; Feld, M. S.; Rava, R. P. *Applied Optics* **1993**, 32, 3585-3595.

(24) Zhang, Q. G.; Muller, M. G.; Wu, J.; Feld, M. S. *Optics Letters* **2000**, 25, 1451-1453.

(25) Bechtel, K. L.; Shih, W. C.; Feld, M. S. *Optics Express* **2008**, 16, 12737-45.

(26) Biswal, N. C.; Gupta, S.; Ghosh, N.; Pradhan, A. *Optics Express* **2003**, 11, 3320-3331.

(27) Shih, W. C.; Bechtel, K. L.; Feld, M. S. *Optics Express* **2008**, 16, 12726-12736.

(28) Reble, C.; Andree, S; Eichler, H.J.; Gersonde, I.; Helfmann, J; *Journal of Biomedical Optics*, **2010**, 15, 037016.

(29) Shao, X.; Bian, X.; Liu, J.; Zhang, M.; Cai, W. *Analytical Methods* **2010**, 2, 1662-1666.

(30) Brereton, R. G.; Lloyd, G. R. *Analyst* **2010**, 135, 230-267.

(31) Hansen, P. C. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*, Society for Industrial and Applied Mathematics; **1998**.

(32) Barman, I.; Kong, C. R.; Dingari, N. C.; Dasari, R. R.; Feld, M. S. *Analytical Chemistry* **2010**, 82(23), 9719-9726.

CHAPTER 6

TISSUE AUTOFLUORESCENCE AND PHOTOBLEACHING

A more generic problem, in comparison to the tissue turbidity phenomenon discussed in the previous two chapters, is the presence of a substantive fluorescence background in acquired Raman spectra. Indeed, despite the recent emergence of Raman spectroscopy as a powerful analytical tool in harsh industrial environments as well as in research settings, the fluorescence arising from either the analyte of interest (if it absorbs at the laser wavelength used for excitation purposes) or, more typically, from adventitious impurities and other interferents present in the sample prevents its broader application. Specifically, biological tissue exhibits intense fluorescence stemming from the complex distribution of intrinsic chromophores. In addition to the broad fluorescence background and the accompanying noise (shot noise and detector noise), the prediction accuracy of a calibration model built for a set of Raman spectra can also be severely compromised by the quenching of the endogenous fluorophores due to the introduction of spurious correlations between analyte concentrations and fluorescence levels. Apparently functional models can be obtained from such correlated samples, which cannot be used successfully for prospective prediction.

In this chapter, we investigate the deleterious effects of photobleaching on prediction accuracy of implicit calibration algorithms, particularly for transcutaneous glucose detection using Raman spectroscopy. Using numerical simulations and experiments on physical tissue models, we show that the prospective prediction error can be substantially larger when the calibration model is developed on a photobleaching correlated dataset as compared to an uncorrelated one. Furthermore, we demonstrate that the application of shifted subtracted Raman

spectroscopy (SSRS) reduces the prediction errors obtained with photobleaching correlated calibration datasets to that obtained with uncorrelated ones.

6.1 Significance

The persistent problem of the fluorescence background in extracting accurate Raman line shapes and consequently in quantifying Raman-active analyte concentrations has been well-documented in the literature [1, 2]. Often, the significantly greater intensity of the sample fluorescence limits analysis to strong Raman bands only [3], unless an active enhancement process is exploited (such as in resonance Raman, surface enhanced Raman, stimulated Raman and coherent anti-Stokes Raman). Moreover, the associated photon shot noise and detector noise, which may have intensities similar to that of the spontaneous Raman signal depending on the signal acquisition times, further compromise the diagnostic capability of the spectroscopic technique.

Given this undesirable situation, investigators in biomedical photonics have gradually shifted towards longer wavelengths in the NIR regime to obtain spectra from specimens. (It should be noted that tissue autofluorescence is one of the reasons for such a shift towards longer wavelengths, the other prominent ones being deeper penetration depth and lower energy optical radiation.) Nevertheless, despite excitation in the NIR regime, the tissue autofluorescence is still several times larger than the signal of the Raman-active analyte of interest.

More importantly, in biological studies, quenching (photobleaching) of the endogenous fluorophores can change the acquired signal appreciably over time. In the context of the following discussion, photobleaching is defined as the reduction in sample autofluorescence intensity over a period of time under sustained/frequent laser exposure. A number of different mechanisms, including deactivation of the excited-state fluorophores upon contact with other

molecules (commonly known as quenchers) in the mixture sample, can contribute to the overall drop in fluorescence intensity [4]. In clinical Raman studies, the biological sample might be subject to a laser beam over a time period varying from a few seconds to several hours. For example, glucose tolerance tests [5] and glucose/insulin clamping studies [6] in humans and animal models can take anywhere between two to six hours. Naturally, fluorescence quenching manifests itself in the Raman spectra acquired over the duration of such studies.

As discussed in Chapter-3, in order to extract quantitative glucose information from the acquired spectra, it is necessary to employ implicit calibration methods such as partial least squares (PLS) and principal component regression (PCR) [7, 8]. These methods require only the concentrations of the analyte of interest and the spectra acquired from the calibration samples to construct a regression model. However, these calibration methods are often misled by chance correlations between non-analyte specific spectral information and the concentrations of the analyte of interest, as originally demonstrated by Arnold *et al.* [9]. Spurious effects such as system drift and covariations among sample constituents can lead to apparently functional models, which cannot be successfully employed for prospective prediction [10].

We hypothesize that sample photobleaching introduces similar spurious correlations in the calibration models for biological Raman spectroscopy, particularly in cases where temporal correlations exist within the concentration dataset. Specifically for the characterization of metabolic activity of the blood glucose regulation system, functional tests, such as glucose loading tests or glucose tolerance tests, are frequently performed where the blood glucose levels are changed by ingestion of glucose or insulin. To prove the clinical feasibility of the spectroscopic techniques (e.g. infrared (IR) and Raman), spectra are typically acquired over the

duration of these studies to form a calibration model (in conjunction with the measured analyte concentrations) for prospective prediction of blood glucose levels. In this chapter, we first assess the robustness (and validity) of models developed on samples undergoing photobleaching during glucose tolerance tests. Using numerical simulations and experimental studies on physical tissue models, we demonstrate that a significant deterioration in prospective prediction accuracy is observed when the model is developed on "correlated samples", i.e. calibration samples that demonstrate statistically significant correlation between (photobleached) fluorescence levels and glucose concentrations, compared to when it is developed on "uncorrelated samples". This is attributed to the construction of a spurious calibration model where the regression vector is partly based on the fluorescence signal rather than only on the analyte of interest (glucose). This result has major implications for diabetics, as any tolerance test based protocol would lead to an approximately monotonic rise in glucose levels over the measurement period due to the inadequate insulin response. Evidently, the resultant covariation between the glucose concentrations and the photobleaching-induced decay in the tissue auto fluorescence levels may lead to systematic errors in the calibration models.

The second goal of this chapter is to examine fluorescence removal methodologies as potential tools for avoiding the pitfalls associated with photobleaching correlated samples. Removal of fluorescence to isolate the Raman bands is a well-studied problem in the vibrational spectroscopy community and several methodologies including time-gating, derivative processing, least squares polynomial subtraction, shifted excitation Raman difference spectroscopy (SERDS) and shifted subtracted Raman spectroscopy (SSRS) have been proposed over the years [11 - 16]. SERDS, which is based on the differential shift response of the Raman and fluorescence signals to the shift in excitation wavelength, offers a promising solution as it

does not suffer from insufficient photon collection problems (as opposed to the time-gating schemes) and does not, in principle, introduce spectral artifacts like the numerical post-processing techniques. Typically, a tunable laser source is employed to obtain the sample spectra at two shifted excitation frequencies, which are then subtracted to eliminate the fluorescence background – leaving only the Raman component as first derivative signals. SSRS is conceptually similar to the SERDS method, except that it involves shifting the spectrograph grating instead of the laser frequency. Although SERDS and SSRS have been employed for reconstruction of Raman spectra from simple mixture models, their efficacy in quantitative problems of concentration prediction particularly when used in conjunction with an implicit calibration method remains largely untested. Here we compare the effectiveness of SSRS to numerical post-processing techniques, namely least squares polynomial subtraction and spectral derivative application, for both photobleaching correlated and uncorrelated samples. We observe that exploiting the shifted Raman spectra in SSRS to reject the fluorescence background reduces the prospective prediction errors for the calibration models developed on the correlated samples to the levels of the models calibrated on uncorrelated samples.

6.2 Experimental methods

6.2.1 Human subject studies

The concentration datasets as well as the tissue fluorescence and photobleaching profile used in our numerical simulations (Sec.-6.3) was originally described in one of our laboratory's previous publications [17] and has also been detailed in Chapter-3. Briefly, NIR Raman spectra were acquired from the forearms of twenty human volunteers undergoing oral glucose tolerance tests (OGTT). An 830-nm diode laser was used for Raman excitation and the back-scattered light

from the tissue was collected on a liquid-nitrogen cooled CCD coupled to a f/1.8i spectrograph. For each volunteer, spectral measurements were initiated following the ingestion of a standard glucose-rich drink. Spectra were acquired for a total of 3 min per data point at 5 min intervals over approximately a two hour time period forming a measurement time series for each volunteer. A representative spectral time series acquired from one of the volunteers is shown in Chapter-3. Figure 6.1 gives the photobleaching profile as measured on this volunteer. It is a curve of I_λ/I_{exc} as a function of time, where I_λ is the area under the spectral curve over a 20 cm^{-1} band surrounding any arbitrarily selected central value (600 cm^{-1} used for data shown here) and I_{exc} is the measured intensity of the excitation source. The measured values of I_λ/I_{exc} were fitted with a double exponential curve [18], given by the red curve in Fig. 6.1.

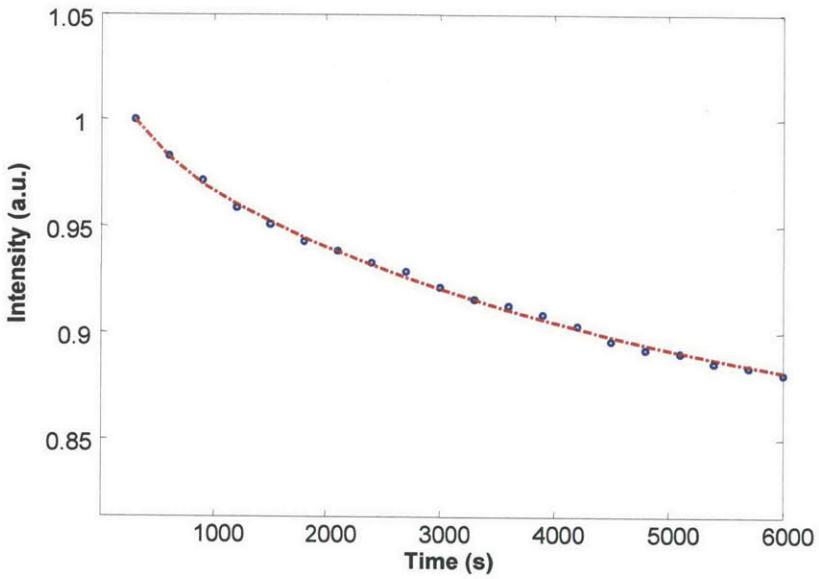


Fig. 6.1 Normalized tissue autofluorescence decay (I_λ/I_{exc}) obtained as a function of time during the OGTT performance (the measured data points are in circles and the dotted line is the best fit double exponential curve).

Concomitant with the spectral measurements, blood samples were also collected using finger sticks at regular 10 min intervals for estimation of glucose levels by a clinical glucose analyzer. For comparison, two such profiles taken from two different volunteers are shown in Fig. 6.2. Profile I shows a rise in blood glucose concentration (to almost double its normal value) followed by a return to original levels – indicative of proper functionality of the blood glucose regulation system in the volunteer. Profile II, in sharp contrast, exhibits a nearly monotonic rise in glucose levels over the test period, characteristic of weak insulin-mediated response in this subject.

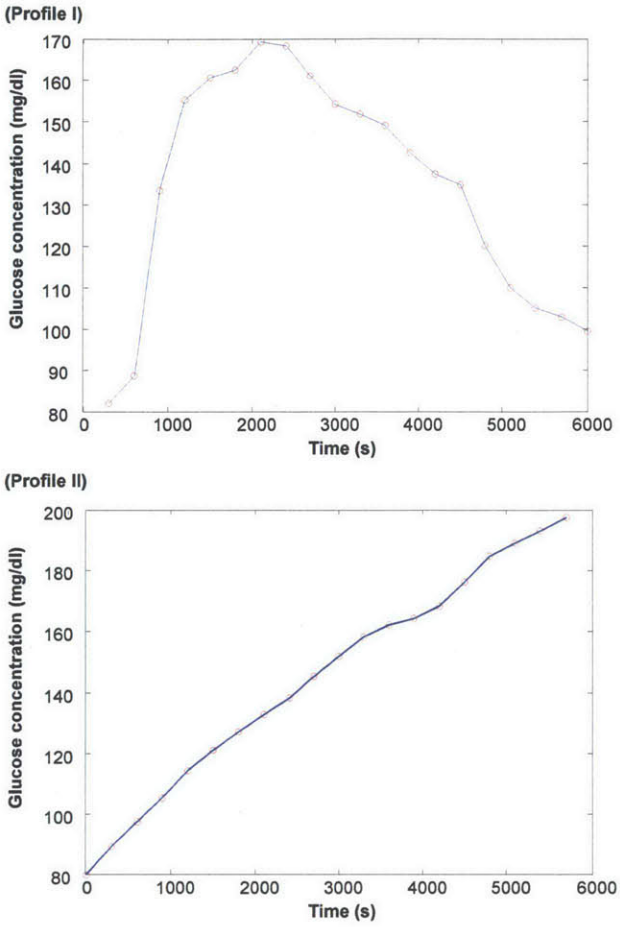


Fig. 6.2 Representative glucose concentration profiles taken from two human volunteers: Profile I (top) and Profile II (bottom).

6.2.2 Tissue phantom studies

Physical tissue models (tissue phantoms) in which the fluorescence and Raman scattering could be precisely varied were used to investigate the prediction performance of calibration models based on the correlated and uncorrelated spectral datasets, respectively. Additionally, in conjunction with the numerical simulations performed in Sec.-6.3, this study was used to understand the effect of fluorescence removal strategies on the prospective prediction accuracy of the developed calibration models.

A total of 50 tissue phantoms were prepared with aqueous solutions of glucose (analyte of interest), urea (spectral interferent) and indocyanine green (ICG). ICG was employed to produce a fluorescence background similar to the tissue autofluorescence observed with near-infrared (NIR) excitation. To mimic the (photobleached) fluorescence correlation observed in glucose tolerance testing of human subjects with impaired glucose tolerance (similar to that observed in Profile II of Fig. 6.2), 20 tissue phantoms were constructed where the correlation between the glucose and ICG concentrations was high ($R^2 \approx 0.95$). The other 30 tissue phantoms had negligible correlation between the glucose and ICG levels ($R^2 \approx 0.1$). For both correlated and uncorrelated tissue phantoms, glucose and ICG concentrations varied from 0.9 to 1.2 M and 0.18 μM to 0.36 μM, respectively. The relatively high concentration of glucose in the tissue phantoms enabled higher fluorescence values to be studied while retaining a satisfactory signal-to-noise ratio for the analyte of interest. The urea concentration in all the tissue phantoms was randomized in the range of 250 to 400 mM.

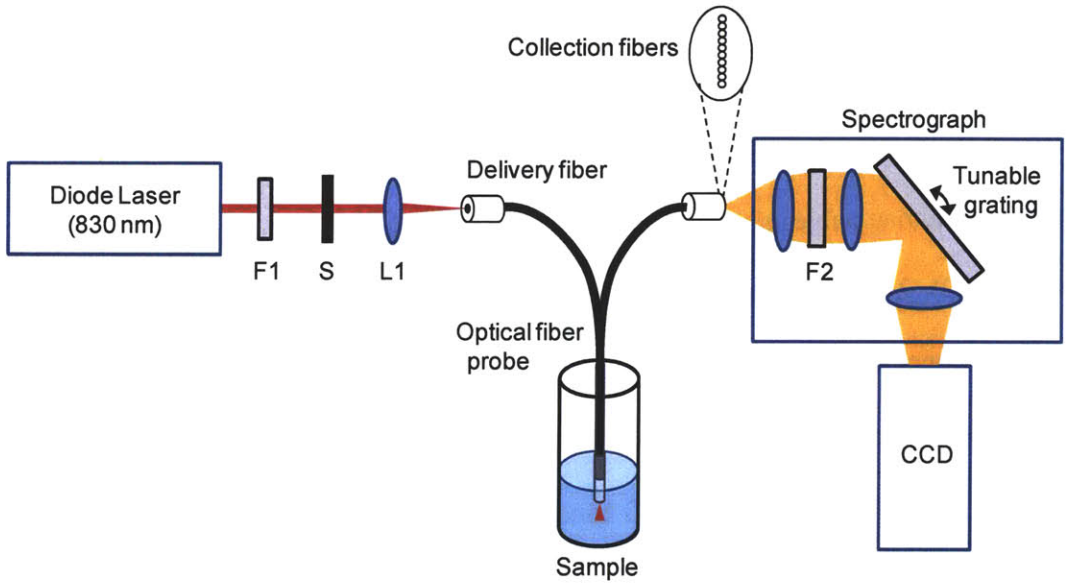


Fig. 6.3 A schematic diagram of the experimental setup. Raman spectra were obtained from tissue phantom solutions using an optical fiber probe, which included a laser light delivery fiber and 10 collection fibers. The spectrograph was equipped with a micrometer, which was used to precisely tune the grating for implementing SSRS. F1: laser line filter, S: shutter, L1: focusing lens for optical fiber coupling, F2: Rayleigh rejection edge filter.

For Raman excitation, an 830-nm diode laser was focused into the delivery fiber (200 μm core diameter) of an optical fiber and filter probe assembly [19] to deliver an average power of ~100 mW on a spot size of ~1 mm^2 at the sample. The sample solution was held in a cylindrical glass vial (1 inch diameter) with a top opening, through which the probe tip was immersed directly into the solution without contacting the container (Fig. 6.3). The back-scattered Raman light was collected by 9 collection fibers (200 μm core diameter) before being dispersed via a tunable f/2.0 spectrograph (LS-785, 5 cm^{-1} resolution, Princeton Instruments) on to a TE-cooled CCD detector (PIXIS 256, Princeton Instruments). The acquisition time for each Raman spectrum was 20 seconds. Spectra of the phantoms were acquired randomly with respect to the

constituents' concentration in order to eliminate any potential temporal correlation. Since the optical fiber probe itself generated prominent Raman spectral features at wavelengths below 800 cm^{-1}, the spectral range of 950 to 1800 cm^{-1} was used for data analysis (PLS calibration and prediction), after vertical binning and cosmic ray removal.

Additionally, in order to investigate strategies for removing the effects of the fluorescence background, SSRS was performed by acquiring the tissue phantom spectra after tuning the spectrograph grating position by ~25 cm^{-1} (approximately the full width at half-maximum of the characteristic glucose peak at 1100 cm^{-1}). The Raman shift axis was calibrated using the standard spectra measured from a 4-acetaminophenol powder sample. The tuning of the grating was monitored by the position of the laser line on the CCD detector such that an 18-pixel shift (corresponding the aforementioned 25 cm^{-1} shift) was maintained between the two positions. Further details of the calibration procedure followed for SSRS implementation are described elsewhere [20]. The difference between the pair of Raman spectra (acquired at two different grating positions) was used for PLS calibration and prediction, and the performance was compared with that of the PLS calibration model constructed using the raw Raman spectra. We have also employed first derivative spectra for comparison with the aforementioned calibration models. For the first derivative spectra, discrete differencing of the acquired spectra was performed. Additionally, we undertook derivative application after smoothing of the acquired spectra (Savitzky-Golay smoothing with 9 point window [21]).

6.3 Numerical simulations

We performed two sets of numerical simulations, using the experimental data from the human subject studies (Sec.-6.2.1), to investigate the questions outlined in Sec-6.1. The first simulation study (Sec-6.3.1) was aimed to assess photobleaching impact on model performance. To

characterize this relationship, we analyzed the robustness and prediction accuracy of calibration models developed on the photobleaching correlated and uncorrelated data sets, respectively. In the second study (Sec-6.3.2), the same analysis was repeated except that the spectra used in the study were processed for fluorescence removal. We investigated two fluorescence removal techniques in our simulations: (a) least squares polynomial subtraction and (b) SSRS. This study was used to compare the improvement in prediction performance obtained in both cases – correlated and uncorrelated – on employment of different fluorescence removal strategies.

6.3.1 Assessment of photobleaching impact on model performance

Simulation calibration spectra were first generated by forming weighted linear combinations of constituent analyte spectra of glucose, creatinine and urea as measured by our Raman system. For our application, glucose represents the analyte of interest, with the other two constituents playing the role of spectral interferents in the mixture model. The assigned weights in the linear combinations correspond to the concentrations of the analytes in the sample at that particular time instance. Two sets of simulation spectra were constructed: (a) a correlated spectral dataset and (b) an uncorrelated spectral dataset. Specifically, the correlation coefficient $\rho_{A,B}$ between two random variables A and B is given by:

$$\rho_{A,B} = \frac{\mathrm{cov}(A,B)}{\sigma_A \sigma_B} \tag{6.1}$$

where cov(A,B) gives the covariance between A and B and σ_A and σ_B are the standard deviations of A and B, respectively [22]. For the considered case, the variables A and B are constituted by the fluorescence intensity values and glucose concentrations, respectively. In the following, we will frequently state the strength of correlation (also called the coefficient of determination) by the R^2 value, which here is the square of $\rho_{A,B}$.

To generate the correlated spectral dataset, the concentrations shown in profile II (Fig. 6.2) were assigned to glucose to formulate the mixture Raman spectra. The glucose concentration values of profile II show a strong negative correlation with the photobleached fluorescence levels plotted in Fig. 6.1 ($\rho_{A,B} \approx -0.95$, $R^2 \approx 0.9$). For the uncorrelated dataset, the concentration values from profile I (Fig. 6.2) were allocated for glucose, as the correlation between this profile and the aforementioned photobleaching curve is insignificant ($R^2 \approx 0.1$). It should be noted that the concentrations of creatinine and urea in both set of simulation spectra were kept nearly constant (i.e. within 5% variation) to replicate the negligible changes observed in the concentration levels of the other analytes during a typical OGTT.

To model the fluorescence we performed hyper-smoothing of a typical tissue spectrum to remove the sharper Raman features and retain only the autofluorescence background. For both correlated and uncorrelated dataset simulations, this hyper-smoothed spectrum was scaled by a factor specified by the photobleaching profile of Fig. 6.1 at each time point before addition to the mixture Raman spectra. Moreover, in order to fully evaluate the impact of fluorescence to Raman amplitude ratio on the performance of the calibration models for both the correlated and uncorrelated cases, the relative intensity of the fluorescence and Raman signals was also varied in the range of 5-20. Finally, noise was generated in accordance with the shot-noise-limited condition (which has previously been observed to hold true for spectra acquired using our Raman system [23]) for addition to the fluorescence-Raman combination spectra. We will revisit the shot-noise limited assumption and its implications in Sec-6.4.2. The signal-to-noise ratio (SNR), defined here as the ratio of the maximum intensity of the fluorescence-Raman combination signal to the mean noise magnitude, was varied from 50 to 200. The range of fluorescence to Raman amplitudes and SNR values mentioned above are consistent with that observed in Raman

spectra acquired for these types of analytical measurements (typical values of observed SNR~100 and fluorescence-to-Raman ratio~10). At each setting of the tunable parameters, i.e. fluorescence to Raman ratio and SNR, 100 calibration spectra were generated for the correlated as well as the uncorrelated datasets.

The prediction spectra were generated following a scheme similar to that stated above except that the concentrations of all the analytes were completely randomized in this dataset. The fluorescence levels in the prediction spectra were also completely randomized to simulate those observed in a set of unknown samples (such as a test set of human volunteers), where the fluorescence levels and the concentrations of the analytes would be expected to possess minimal correlation with one another. This situation typically occurs when a calibration model developed during a tolerance or loading test is used prospectively on different subjects to determine their glucose levels. For a given setting of the fluorescence to Raman ratio and SNR level, 50 such spectra were included in the prediction data set.

Partial least squares (PLS) analysis was used to develop the calibration models based on leave-one-out cross-validation on the calibration (training) set. The reported final errors, root-mean-square error of prediction (RMSEP), were obtained by employing these calibration models on the independent prediction sets. To ensure reproducibility of our results, the entire calibration-prediction procedure was repeated 20 times.

6.3.2 Assessment of fluorescence removal methods

In this simulation study, we followed a similar algorithm as that outlined in Sec-6.3.1, except that all the spectra (i.e. calibration and prediction, both for correlated and uncorrelated cases) were processed for fluorescence removal. Here, we employed SSRS to quantify its benefits,

relative to that of numerical post-processing schemes, for both the photobleaching correlated and uncorrelated cases.

For our simulations, we generated a second set of calibration and prediction spectra for both the correlated and uncorrelated datasets by shifting the composite mixture spectra (Raman plus fluorescence) by 25 cm^{-1}. These shifted spectra were then subtracted from the original set of spectra to obtain the Raman difference spectra (plus shot noise and residual fluorescence background). While most of the previous research has primarily been focused on reconstruction of the intrinsic Raman spectrum from the difference spectra [24 - 26], we have employed the difference spectra directly in analysis to extract concentration information through PLS. The direct incorporation of the difference spectra into calibration analysis is advantageous as it avoids the artificial spectral features and noise (ringing behavior) that may appear due to the application of a reconstruction method (such as Fourier deconvolution).

For fair comparison with numerical strategies, we implement three variations of least- squares polynomial subtraction as alternative fluorescence removal tools. First, a lower order polynomial is fit to the broad fluorescence background and subtracted. Specifically, we employ a second or third order polynomial for the PLS analysis, depending on which order provides the least error in cross-validation. This approach is widely pursued for signals with broad fluorescence backgrounds because higher order polynomials tend to overfit the data, as also noted by Beier *et al.* [27]. Second, an algorithm developed by Mahadevan-Jansen and co-workers [13], which automatically fits a modified polynomial just below the original spectrum, is used. Third, a recently developed adaptive 'minmax' scheme [28] is used to subtract the background fluorescence. PLS analysis is undertaken in conjunction with each of the three numerical processing schemes.

6.4 Results and discussion

6.4.1 Numerical simulations

6.4.1.1 Impact of photobleaching on model performance

As detailed in Sec-6.3.1, PLS was first used in a leave-one-out cross validation routine on the calibration spectral set. The number of loading factors used to build the final model for prediction was in accordance with the number of constituents that provided the minimum error in cross-validation. The results of the simulations at varying levels of Raman-to-noise ratio for the photobleaching correlated and uncorrelated cases are plotted in Fig. 6.4, with error bars indicating the standard deviation of the RMSEP values at each setting. It is evident that the calibration model developed on the correlated calibration spectra always performs worse than the model developed on the uncorrelated set, irrespective of the SNR value or the fluorescence-to-Raman amplitude ratio. For example, at a setting of fluorescence-to-Raman of 10 and SNR of 100 (values observed typically for tissue Raman spectra), the mean prediction errors were measured to be 16.3 mM and 5.3 mM for the correlated and uncorrelated cases, respectively, thereby demonstrating an approximately three-fold increase in the prediction error. Expectedly, for lower levels of Raman-to-noise ratio the prediction performance of both models drop considerably. (It is noted that changing the fluorescence-to-Raman ratio or the SNR has an identical effect, as they both modulate the Raman-to-noise ratio.) Importantly, the performance of the calibration model built on the correlated dataset drops off much more significantly than that developed on the uncorrelated dataset – as evidenced by the widening gap between the bars at higher fluorescence-to-Raman (i.e. lower Raman-to-noise) settings. This can be attributed to the fact that it becomes more difficult to distinguish the Raman spectrum of glucose from the confounding fluorescence background in a noisier spectral set. The simplest experimental

solution to reducing the disparity between the correlated and uncorrelated datasets is to increase the overall SNR, which would then decrease the RMSEP - however this would come at the cost of either increased exposure time or increased laser power.

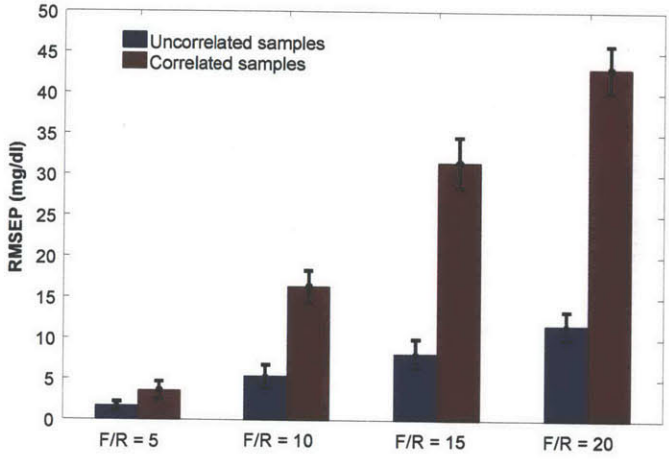


Fig. 6.4 Bar plot of RMSEP obtained for calibration models developed on photobleaching correlated and uncorrelated datasets, respectively, as a function of increasing fluorescence-to-Raman ratio (i.e. decreasing Raman-to-noise ratio). Here, the SNR is held constant (100). Identical results are obtained by changing the SNR while holding the fluorescence-to-Raman ratio fixed.

The comparison of PLS prediction performance between correlated and uncorrelated datasets (Fig. 6.4) can be further generalized to show that the amount of deterioration in the prediction performance is dependent on the magnitude of correlation (R^2) between glucose concentration and tissue fluorescence. In Fig. 6.5, RMSEP increases with increasing R^2 for all values of fluorescence-to-Raman ratios. At low values of R^2, the RMSEP appears to be independent of R^2, but displays a steep rise after a threshold value of R^2. This suggests that the prediction error at low values of R^2 is determined primarily by the spectroscopic SNR, while at higher values of R^2,

the prediction error depends on both SNR and R^2. This can also be visualized by observing the widening gap between any pair of RMSEP curves with increase in the R^2 value in Fig. 6.5.

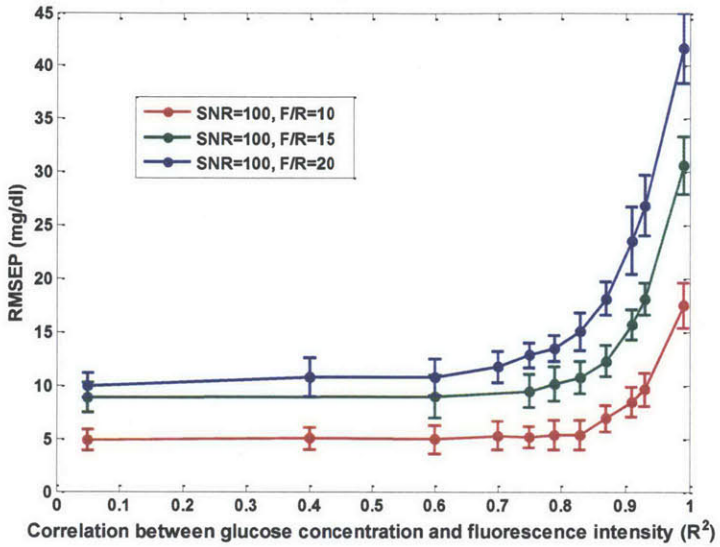


Fig. 6.5 RMSEP of simulations as a function of the correlation (R^2) between glucose concentration and fluorescence intensity. The fluorescence-to-Raman ratio (F/R) was varied from 10, 15 to 20, and 20 simulations were performed for each case.

The results of Fig. 6.4-6.5 have important clinical implications, particularly for the correlated case. Such a situation is representative of experimental observations from OGTTs performed on diabetic and insulin-resistant patients. Since it is difficult to perform glucose/insulin clamping studies in normal and, particularly, diabetic populations, tolerance tests have remained the most viable protocol for the development of a calibration model using spectroscopic techniques [29, 30]. Moreover, while test strip based measurements and glycosylated hemoglobin (HbA1c) examination are more extensively used at the present time, tolerance tests remain the gold standard of clinical tests of insulin/glucose control system and have been used on a regular basis to screen for gestational diabetes [31, 32] and to determine the level of insulin resistance [33].

However, our results indicate that photobleaching introduces a systematic error in Raman spectroscopic calibration model development using tolerance tests, particularly for people with impaired glucose regulation. Put another way, any model developed on a diabetic population is susceptible to substantial errors in prospective prediction, even when used for diabetic patients only. This is clearly undesirable. The potential remedies to break the unwanted correlation include a change in experimental study design (e.g. point-of-care measurements across a large population) and explicit removal of the source of such correlation, i.e. the varying fluorescence background. The inclusion of a large number of human volunteers in the calibration set may, however, introduce additional complexities such as large variations in tissue optical properties and skin heterogeneity.

6.4.1.2 Relative performance of fluorescence removal methods

For this study, the fluorescence removed spectra – SSRS-processed and least squares polynomial subtracted – were used in conjunction with the concentration datasets to build the calibration models (as detailed in Sec-6.3.2). The results of prospective prediction of the five models (i.e. based on raw, lower order polynomial subtracted, modified polynomial subtracted, minmax fit subtracted and SSRS processed spectra) are summarized in Fig. 6.6. For this set of simulations, experimentally observed values of SNR (of 100) and fluorescence-to-Raman ratio (of 10) were used.

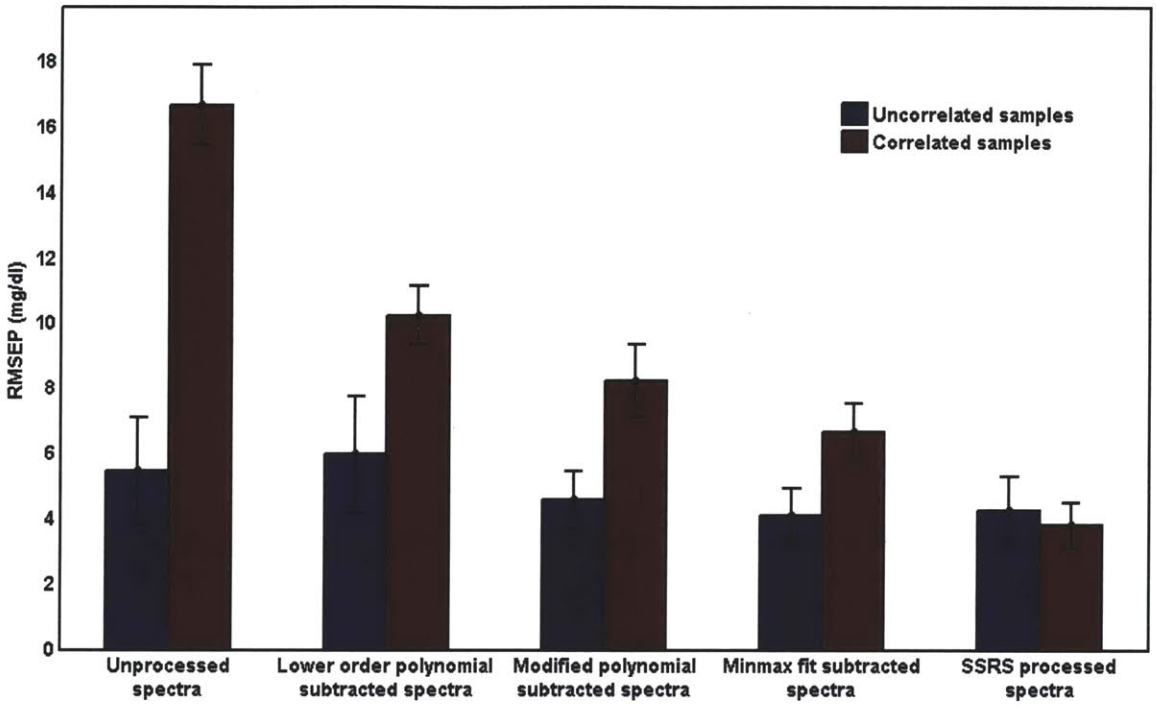


Fig. 6.6 Bar plot of RMSEP values obtained for glucose concentrations from simulations on photobleaching correlated (red) and uncorrelated (blue) datasets. The groups represent calibration and prediction performed using the following types of spectra: (from left to right) unprocessed, lower order polynomial subtracted, modified polynomial subtracted, minmax fit subtracted and SSRS processed spectra, respectively.

For the correlated case, SSRS provides substantial reduction of prediction error (from 16.7 to 3.8 mg/dl), even lower than the levels obtained using the raw uncorrelated calibration model (5.5 mg/dl). Conventional lower-order polynomial subtraction, on the other hand, provides only about half the amount of improvement (from 16.7 to 10.3 mg/dl) as SSRS processing for the correlated case. This disparity in results could be attributed primarily to the introduction of non-physical negative regions in the spectra and associated artifacts due to the process of fluorescence removal using the polynomial subtraction technique. Newer variants of the polynomial subtraction technique, namely the modified polynomial (8.3 mg/dl) and adaptive minmax fitting

(6.7 mg/dl), provide lower average RMSEP values than conventional polynomial subtraction. Clearly, the inclusion of meaningful constraints in these techniques enhances the prospective prediction accuracy, though not to the levels obtained using SSRS.

Our results for the uncorrelated case suggest that application of SSRS or numerical processing do not cause similar changes in prediction performance from the raw spectra results. This is not unexpected because implicit calibration schemes (PLS) do a reasonable job of separating the analyte of interest (e.g. glucose) from the confounding analytes (including tissue fluorescence), especially when reasonable levels of SNR are achieved [34].

Finally, it is important to understand that our numerical simulations have been based on the assumption that the main factor limiting SNR is the shot noise in the system. Although this assumption holds good for experiments on tissue samples where exposure times are limited by other constraints, this may not always be the case (even for glucose tolerance studies, where one can afford longer acquisition times due to the relatively slow physiological responses). As has been pointed out by Bell and co-workers [20, 35, 36], there are many situations where it is observed that the random but fixed variations in pixel-to-pixel response on the CCD might actually be more dominant than shot noise. This is especially true when high signal levels can be achieved and exposure times can be increased without any significant downside (such as when chemical mixtures and powder samples are tested). SSRS will clearly provide a bigger improvement in such applications because the major cause of noise arising from the irregularity in detector response can be canceled when the shifted spectra are subtracted.

6.4.2 Experimental studies on tissue phantoms

To validate our simulation results, we analyzed the spectral datasets obtained in our tissue phantom studies. Figure 6.7 shows: (A) a raw Raman spectrum obtained from a representative tissue phantom, (B) SSRS data obtained by subtracting two raw spectra, obtained at slightly different spectrograph grating positions, 25 cm^{-1} apart and (C) numerical first derivative of the raw spectrum shown in (A). It is clear that the fluorescence background can be suppressed using either shifted subtraction at two grating positions (SSRS) or taking the first derivative. However, the SNR of the SSRS data appears to be significantly better than that of the first derivative spectra. This can be attributed to the cancellation of the fixed pattern noise on the CCD detector, as discussed above. It should be noted that having the correct amount of shift between the two acquired spectra is critical to the application of SSRS as too small a shift would result in poor SNR and too large a shift would cause improper background suppression. For the SSRS spectrum, we also observe the presence of a small residual background because of the corresponding shift in fluorescence in the two sets of spectra.

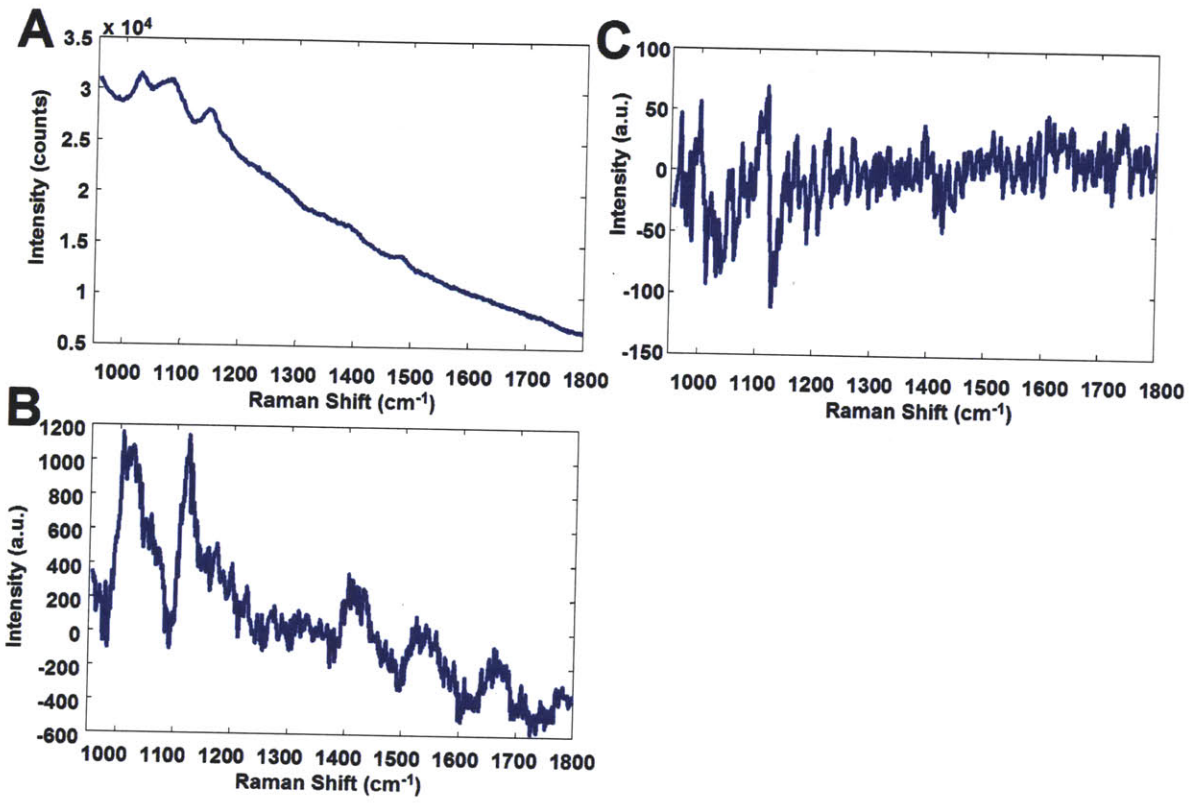


Fig.6.7 (A) Representative Raman spectrum acquired from a tissue phantom; (B) SSRS spectrum obtained by subtracting two spectra, obtained at spectrograph grating positions 25 cm^{-1} apart; (C) First derivative spectrum.

We first created the prediction dataset by randomly extracting 10 uncorrelated tissue phantom data. PLS calibration models were then constructed based on the raw spectra, SSRS spectra and first derivative spectra for both the fluorescence-correlated and uncorrelated datasets (which consisted of 20 tissue phantoms each). Model performance during optimization was judged by comparing the standard error for leave-one-out cross-validation in the calibration data set. The prediction set was withheld from all optimizations and was solely employed to evaluate the performance of the final optimized models. The whole procedure was iterated fifty times to

obtain an average and standard deviation of the RMSEP values. Table 6.1 summarizes the results of prospective prediction.

Table 6.1: Summary of mean and standard deviation (in parentheses) of the RMSEP values obtained from the tissue phantom experiments for both the correlated and uncorrelated datasets.

	Raw	SSRS	1st derivative	1st derivative after smoothing
Correlated	53.8 (±7.8) mM	42.7 (±7.3) mM	66.3 (±9.2) mM	56.0 (±8.1) mM
Uncorrelated	39.1 (±6.7) mM	42.4 (±6.2) mM	65.0 (±8.1) mM	55.5 (±8.9) mM

We observe that the prediction performance of the calibration model built on the fluorescence correlated dataset is substantially worse than that built on the uncorrelated phantoms. The mean prospective prediction errors are 53.8 mM and 39.1 mM for the correlated and uncorrelated calibration models, respectively. Importantly, we find that the application of SSRS reduces the error in the correlated calibration model from 53.8 mM to 42.7 mM to the levels achieved with the uncorrelated dataset. We also observe that the application of SSRS marginally increases the prediction error of the uncorrelated dataset from 39.1 mM to 42.4 mM. This may be attributed to the imperfect rejection of the fluorescence background by SSRS application. Nevertheless, the two primary results observed in our previous numerical simulations are reinforced by this tissue phantom study: (a) calibrations developed from the fluorescence-correlated datasets have a greater prediction error than calibrations developed with fluorescence-uncorrelated datasets and (b) the application of frequency shifted spectra provides more accurate concentration predictions than that obtained from the unprocessed spectra, particularly for the fluorescence-correlated dataset.

In sharp contrast to SSRS, the first derivative application appears to adversely impact the prediction performance for both the fluorescence correlated and uncorrelated datasets. One would expect that this is an outcome of the reduction in SNR of the first derivative spectra observed in Fig. 6.7(c). This is consistent with previous observations, especially when an additional filtering procedure is not employed [2]. It has also been previously reported that the derivative method makes the lower intensity Raman bands difficult to identify in the processed spectrum [13]. Smoothing of the acquired spectra before the differentiation procedure, however, reduces the average RMSEP value, probably due to the resultant increase in SNR [12], compared to simple derivative application.

Taken in conjunction with the least squares polynomial subtraction results from our simulations, these results indicate that while the numerical fluorescence removal strategies are able to suppress the fluorescence background, the instrumentation-based SSRS technique provides better quantitative predictions. Furthermore, the tissue phantom studies verify the robustness of the SSRS methodology in a realistic setting.

6.5 Summary

Development of a robust calibration model for transcutaneous blood analyte detection using Raman spectroscopy may be severely impeded by the presence of photobleaching-induced varying fluorescence background. In this chapter, we have shown that significant deterioration in model performance can be observed due to the photobleaching induced spurious correlations. At typically observed levels of SNR and fluorescence to Raman ratio, this may result in a factor of three increase in prediction error. This has serious implications in developing a calibration model, particularly for diabetic patients using glucose tolerance test-like protocols. To remedy the situation, one needs to incorporate some combination of a change in experimental study

design and fluorescence removal strategy. Evidently, a similar photobleaching-induced problem can be expected in other biological and industrial applications, especially when the measurement period is of similar time scales as the sample fluorescence life time.

Furthermore, we have demonstrated that SSRS reduces the prospective prediction error of the developed calibration models by rendering the model less susceptible to spurious effects associated with photobleaching correlated samples. In this regard, SSRS appears to substantially outperform conventional numerical post-processing schemes such as least squares polynomial subtraction and derivative application. It is worth noting that newer numerical processing schemes, such as the adaptive minmax approach and the automated background removal algorithm proposed by Berger and co-workers [27], may provide better results than the conventional numerical schemes and the advantages of SSRS application over these methods need to carefully studied for the specific application. Nevertheless, our results imply that the problem of photobleaching-induced spurious correlation can be better solved by incorporating a tunable spectrograph, as compared to numerical processing schemes. Such implementation ensures that the principal advantages of glucose tolerance tests (i.e. the convenience of testing and the lack of stringent glucose and insulin control requirements as are necessary for clamping studies) are not sacrificed.

Acknowledgements

Dr. Gajendra Pratap Singh helped with the problem identification. In addition, Professor Peter So's course on *Optical microscopy and Spectroscopy for Biology and Medicine* aided in demarcating the "physically plausible" improvements from the "theoretically impossible" ones. Many of the figures and text were reproduced from the journal article [37] with permission from SPIE.

6.6 References

(1) Chase, D. B. *Journal of American Chemical Society*, 1986, 108, 7485-7488.

(2) Lewis, I. R.; Edwards, H. G. M. *Handbook of Raman spectroscopy: from the research laboratory to the process line*, Marcel Dekker, Inc., New York, **2001**.

(3) Matousek, P.; Towrie. M.; Parker, A. W. *Journal of Raman Spectroscopy*, **2002**, 33(4), 238-242.

(4) Lakowicz, J. R. *Principles of fluorescence spectroscopy*, Springer, **1999**.

(5) American Diabetes Association, *Diabetes Care*, **2009**, 32, S13-S61.

(6) DeFronzo, R. A.; Tobin, J. D.; Andres, R. *Am. J. Physiol. Gastrointest. Liver Physiol.*, **1979**, 237(3), G214-G233.

(7). Wold, S.: Martin, H.; Wold, H. *Lecture Notes in Mathematics*, Springer-Verlag, Heidelberg, **1983**.

(8) Gunst, R. F.; Mason, R. L. *Regression Analysis and its Applications*, Marcel Dekker Inc., New York, **1980**.

(9) Arnold, M. A.; Burmeister, J. J.; Small, G. W. *Analytical Chemistry*, **1998**, 70(9), 1773-1781.

(10) Liu, R.; Chen, W.; Gu, X.; Wang. R. K.; Xu, K *J. Phys. D: Appl. Phys.*, **2005**, 38, 2675–2681.

(11) Yaney, P. P. *J. Opt. Soc. Am.*, **1972**, 62(11), 1297-1303.

(12) Mosier-Boss, P. A.; Lieberman, S. H.; Newbery, R.; *Applied Spectroscopy*, **1995**, 49(5), 9.

(13) Lieber, C. A.; Mahadevan-Jansen, A. *Applied Spectroscopy*, **2003,** 57(11), 1363-1367.

(14) Shreve, A. P.; Cherepy, N. J.; Mathies, R. A. *Applied Spectroscopy*, **1992**, 46(4), 5.

(15) Brennan III, J. F. B.; Wang, Y.; Dasari, R. R.; Feld, M. S. *Applied Spectroscopy*, **1997**, 51(2), 201-208.

(16) Morris, M. D.; Matousek, P.; Towrie, M.; Parker, A. W.; Goodship, A. E.; Draper, E. R. C. *Journal of Biomedical Optics*, **2005**, 10(1), 014014.

(17) Enejder, A. M. K.; Scecina, T. G.; Oh, J.; Hunter, M.; Shih, W.-C.; Sasic, S.; Horowitz, G.; Feld, M. S. *Journal of Biomedical Optics* **2005**, 10, 031114.

(18) Zeng, H.; MacAulay, C.; Palcic, B.; McLean, D. I. *Proceedings of SPIE, 1882(Laser-Tissue Interaction IV)*, **1993**, 278-290.

(19) Ščepanovic, O.; Volynskaya, Z.; Kong, C. R.; Galindo, L.; Dasari, R. R.; Feld, M. S. *Review of scientific instruments*, **2009**, 80(4), 043103.

(20) Bell, S. E. J.; Bourguignon, E.; Dennis, A. *Analyst*, **1998**, 123, 1729–1734.

(21) Savitzky, A.; Golay, M. J. E. *Analytical Chemistry*, **1964**, 36(8), 1627–1639.

(22) Brereton, R. G. *Applied Chemometrics for Scientists*, John Wiley & Sons Ltd.,

Chichester, West Sussex, England; **2007**.

(23) Shih, W. C. *Quantitative biological Raman spectroscopy for non-invasive blood analysis*, Massachusetts Institute of Technology, Mechanical Engineering, Ph.D. Thesis, Cambridge, MA, **2007.**

(24) Zhao, J.; Carrabba, M. M.; Allen, F. S. *Applied Spectroscopy*, **2002**, 56(7), 12.

(25) Osticioli, I.; Zoppi, A.; Castellucci, E. M. *Applied Spectroscopy*, **2007**, 61(8), 839-844.

(26) McCain, S. T.; Willet, R. M.; Brady, D. J. *Optics Express*, **2008**, 16(15), 10975-10991.

(27) Beier, B. D.; Berger, A. J. *Analyst*, **2009**, 134, 1198-1202.

(28) Cao, A.; Pandya, A. K.; Serhatkulu, G. K.; Weber, R. E.; Dai, H.; Thakur, J. S.; Naik, V. M.; Naik, R.; Auner, G. W.; Rabah, R.; Freeman, D. C. *Journal of Raman spectroscopy*, **2007**, 38(9), 1199 - 1205.

(29) Heise, H. M.; Marbach, R.; Koschinsky, T. H.; Gries, F. A. *Artificial Organs*, **1994**, 18(6), 439-447.

(30) Fischbacher, C.; Jagemann, K. U.; Danzer, K.; Müller, U. A.; Papenkordt, L.; Schüler, J. *Fresenius Journal of Analytical Chemistry*, **1997**, 359(1), 78-82.

(31) O'Sullivan, J. B.; Mahan, C. M. *Diabetes*, **1964**, 13, 278-285.

(32) Sacks, D. A.; Greenspoon, J. S.; Abu-Fadil, S.; Henry, H. M.; Wolde-Tsadik, G.; Yao, J. F. F. *Am. J. Obstet. Gynecol*, **1995**, 172(2), 607-614 .

(33) Phillips, D. I.; Clark, P. M.; Hales, C. N.; Osmond, C. *Diabetic Med.*, **1994**, 11(3), 286-92

(34) Mazilu, M.; Luca, A. C. D.; Riches, A.; Herrington, C. S.; Dholakia, K. *Optics Express*, **2010**, 18(11), 11382-11395.

(35) Bell, S. E. J.; Bourguignon, E. S. O.; Dennis, A. C.; Fields, J. A.; McGarvey, J. J.; Seddon, K. R. *Analytical Chemistry*, **2000**, 72(1), 234-239.

(36) O'Grady, A.; Dennis, A. C.; Denvir, D.; McGarvey, J. J.; Bell, S. E. J. *Analytical Chemistry*, **2001**, 73(9), 2058-2065.

(37) Barman, I.; Kong, C. R.; Singh, G. P.; Dasari, R. R. *Journal of Biomedical Optics*, **2011**, 16(1), 011003.

CHAPTER 7

DYNAMIC CONCENTRATION CORRECTION (DCC)

Chapters 4 through 6 detail some of the technical (spectroscopic) challenges associated with validating a non-invasive glucose sensor and our proposed solutions to tackle these challenges. Nevertheless, the complex nature of any minimally invasive or non-invasive glucose sensing approach introduces issues having a physiological/anatomical character. In other words, these challenges relate to the human physiology, rather than specifically to the technology involved in the diagnostic approach. Such challenges, some of which are specific to the anatomical measurement site, include physiological fluctuations arising from variable blood flow rates in an individual, skin heterogeneity within the same anatomical zone and even motional artifacts. Most of these challenges can be reduced (if not completely eliminated) by appropriate selection of measurement site (such as in the earlobe, inner lip and thennar fold) and by development of a suitable mechanical interface that ensure site reproducibility, as discussed in Chapter-9. However, a significantly bigger problem originates from the physiological glucose dynamics, namely the presence of differential levels of glucose in the blood and interstitial fluid (ISF) compartments. This is a particularly important problem for spectroscopic techniques, which predominantly probe ISF glucose, creating inconsistencies in calibration, where blood glucose measurements are used as reference.

To overcome this problem, we present a dynamic concentration correction (DCC) scheme, based on the mass transfer of glucose between blood and ISF, to ensure consistency with the spectral measurements. The proposed formalism allows the transformation of glucose in the concentration domain, ensuring consistency with the acquired spectra in the calibration model.

We demonstrate that the predicted glucose concentrations using DCC-based calibration model closely match the measured glucose concentrations, while those generated with the conventional calibration methods show significantly larger deviations from the measured values. In addition, we provide an analytical formula for a previously unidentified source of limiting uncertainty arising in spectroscopic glucose monitoring from a lack of knowledge of glucose kinetics in prediction samples. A study with human volunteers undergoing glucose tolerance tests indicate that this lag uncertainty, which is comparable in magnitude to the uncertainty arising from systematic noise and non-orthogonality in the spectral dataset, can be reduced substantially by employing the DCC in spectroscopic calibration.

7.1 Significance

The phenomenon of physiological lag between the blood and ISF glucose levels is well-known in the medical community [1-4] (aptly defined as "a tale of two compartments" by Cengiz and Tamborlane in their review paper on this subject [5]). In this work, we will primarily focus on the blood-ISF glucose relationship in the dermal-subcutaneous tissue, as it is the area of interest for all minimally invasive and non-invasive glucose sensing approaches.

Briefly, glucose diffuses from the capillary endothelium to the ISF space across an established concentration gradient without the aid of an active transporter [6]. The amount of glucose transported to a specific anatomical region depends on the blood flow to the area. On the other hand, ISF glucose values are determined by the rate of glucose diffusion between the two compartments and the glucose consumption rate by subcutaneous tissue cells. In addition, the ISF glucose level can also be said to be a function of the metabolic rate of the adjacent cells, insulin activity, and the permeability of the capillary that is itself affected by nerve stimulation

[7]. Taken together, this forms a complex functional relationship between the glucose levels in the two compartments (loosely expressed as "physiological lag"), which is accentuated during times of rapid changes of glucose levels. (Strictly speaking, there does not exist a single-valued relation between the two values as described in the following sections.) Specifically, this lag time creates an inconsistency in spectroscopic calibration algorithms, which are based on reference blood glucose concentrations and the acquired tissue spectra. This inconsistency arises from the fact that the spectroscopic techniques primarily probe ISF glucose [4], due to the relatively shallow penetration depth (~500μm-1mm) of near-infrared (NIR) light in tissue and the small density of blood vessels in the superficial layers of the skin [8].

The inconsistency in calibration presents a severe hindrance not only for spectroscopy based non-invasive glucose monitoring but also for present day minimally invasive electrochemical sensors, which base their glucose estimates on interstitial fluid measurements. Indeed, as pointed out in [5], the physiological lag introduces systematic errors during calibration which adversely impact long-term sensor performance, even in the presence of a positive correlation between blood and ISF glucose [9-11]. Such diagnostic errors may lead to unnecessary insulin bolus, which significantly increases the risk of hypoglycemia [12, 13]. The presence of systematic errors is one of the reasons that such continuous glucose monitoring sensors need to be re-calibrated against fingerstick measurements at regular intervals.

Put another way, for all ISF glucose-based sensors, spectroscopic or otherwise, the underlying assumption is that the blood-to-ISF glucose gradient remains constant over the measurement range [3]. However, this assumption fails if the sensor is calibrated during rapid changes in blood glucose, as are encountered during glucose tolerance tests, which provide the

most viable protocol for the development of calibration models using spectroscopic techniques [14]. Calibration during such non-equilibrium conditions leads to large errors in the developed model. To account for the differences in blood and ISF glucose, Bonnecaze and co-workers established the first substantive models for ISF glucose by considering the ISF and blood glucose to reside in two compartments, and performing a mass balance between them [15, 16]. Using amperometric glucose sensors implanted in rats, they demonstrated that accurate estimates of blood glucose concentration can be extracted from subcutaneous ISF glucose-based measurements. However, no analogous models exist for development of spectroscopic calibration algorithms, which are inherently more complex because of the multivariate nature of the data.

Furthermore, even if an accurate calibration model can be established by performing all measurements under equilibrium conditions (e.g. by employing glucose clamps), the lack of knowledge of glucose kinetics in prediction samples would introduce an uncertainty in the concentration estimates. Such prediction uncertainties may lead to inappropriate treatments too. Previously several research groups, including our own laboratory, have assessed uncertainty in spectroscopic prediction based on the noise in the spectral and concentration datasets and the non-orthogonality (spectral overlap) of the analyte of interest with the other sample constituents [17-19]. However, the uncertainty introduced due to the physiological lag in the prediction samples remains unexplored.

This chapter presents a new spectroscopic calibration scheme based on "dynamic concentration correction" (DCC), which is based on a two-compartment mass transfer formulation of blood and ISF glucose and is designed to provide an accurate estimate of glucose

concentrations for non-invasive measurements. These transformations are performed iteratively in conjunction with an implicit calibration method, such as partial least squares (PLS), to form an accurate and consistent regression model. The resulting calibration model can be used on a new set of acquired spectral samples – "the prediction set" - to calculate the ISF glucose concentrations of the samples. Subsequent application of the DCC model converts the estimated ISF glucose concentrations to the equivalent blood glucose concentrations of the prediction samples.

This work employs Raman spectroscopy as a specific example to demonstrate the effectiveness of the new calibration method, with the understanding that this scheme can be similarly applied to other spectroscopic techniques, such as NIR absorption. Using blood and ISF glucose concentration datasets obtained by Steil *et al.* [20], we first demonstrate that predicted glucose concentrations using the DCC calibration model closely match the measured blood glucose concentrations, whereas those generated solely by the conventional implicit calibration methods show significantly larger deviations from the measured values. These results are further validated on spectral and concentration datasets obtained from our laboratory's clinical studies on human volunteers undergoing glucose tolerance tests.

In addition, we derive analytical expressions for the limiting uncertainty introduced in the concentration predictions due to presence of the physiological lag – with and without application of DCC. Here, limiting uncertainty is defined as the uncertainty in concentration estimate in the case where all modeling noise is disregarded, i.e. where the calibration model is assumed to be completely accurate and noise free. Employing the human volunteer data, we find that the concentration uncertainty due to the lag phenomenon is comparable to that arising from the noise and overlap in the prediction spectra and that this major source of uncertainty can be

significantly reduced (approximately six-fold) when DCC is used, providing further motivation for its use in spectroscopy-based transcutaneous blood glucose monitoring.

7.2 Theoretical formulation

The primary motivation for proposing a new spectroscopic calibration method for blood glucose detection is to establish consistency in the calibration model, which maps the spectral measurements to the glucose concentrations. The conventional linear calibration equation can be written as [21]:

$$\mathbf{b} = \mathbf{S}^{*}\,\mathbf{c}, \tag{7.1}$$

where **b** is the spectrum of regression coefficients (also called the regression vector), **S** is the matrix of calibration spectra and **c** is the vector of measured concentrations of the analyte of interest in the calibration samples. $\mathbf{S}^{*}$ is the appropriate inverse of **S**, as evaluated by the calibration method of choice. (Lowercase boldface represents a vector and uppercase boldface denotes a matrix.)

As mentioned in 7.1, since the spectral measurements are predominantly contributed by ISF glucose, the relevant input concentrations to the implicit calibration method should incorporate the ISF glucose concentrations. However, the ISF glucose concentrations are typically not available in a real-life clinical setting – instead, blood glucose values obtained from frequent blood withdrawals are used as reference concentrations. This creates a regression vector, which is neither based completely on blood glucose nor on ISF glucose, but a mixture of the two contributions. This, in turn, also creates a problem in the prediction step, where the predicted glucose concentration c is obtained by a scalar product of the regression vector **b** and the spectrum acquired from the prediction sample **s**:

$$c = \mathbf{s}^{T} \cdot \mathbf{b}. \tag{7.2}$$

(Lowercase italics indicate a scalar quantity and the superscript T denotes the transpose of the vector.) In the conventional calibration framework, the predicted glucose concentration is reported as the blood glucose concentration, although this clearly is not an accurate representation.

To correct for this discrepancy, we propose a new calibration methodology (DCC) in which the concentrations are appropriately changed to conform to the spectral measurements. The transformation in the concentration domain is based on a two-compartment mass transfer model, which establishes a well-defined relationship between blood glucose, ISF glucose and the system parameters. Specifically, we perform the following two transformations in DCC:

(a) Pre-calibration DCC (PC-DCC): Transform the blood glucose concentrations in the calibration dataset to their corresponding ISF values before inputting into the implicit calibration method. This ensures that the regression vector is solely based on ISF glucose contributions.

(b) Post-prediction DCC (PP-DCC): Re-transform the predicted ISF glucose concentration, which is determined by Eq. (2), to the corresponding blood glucose value.

The conceptual differences between the conventional and proposed (DCC-based) calibration methods are shown in Fig. 7.1.

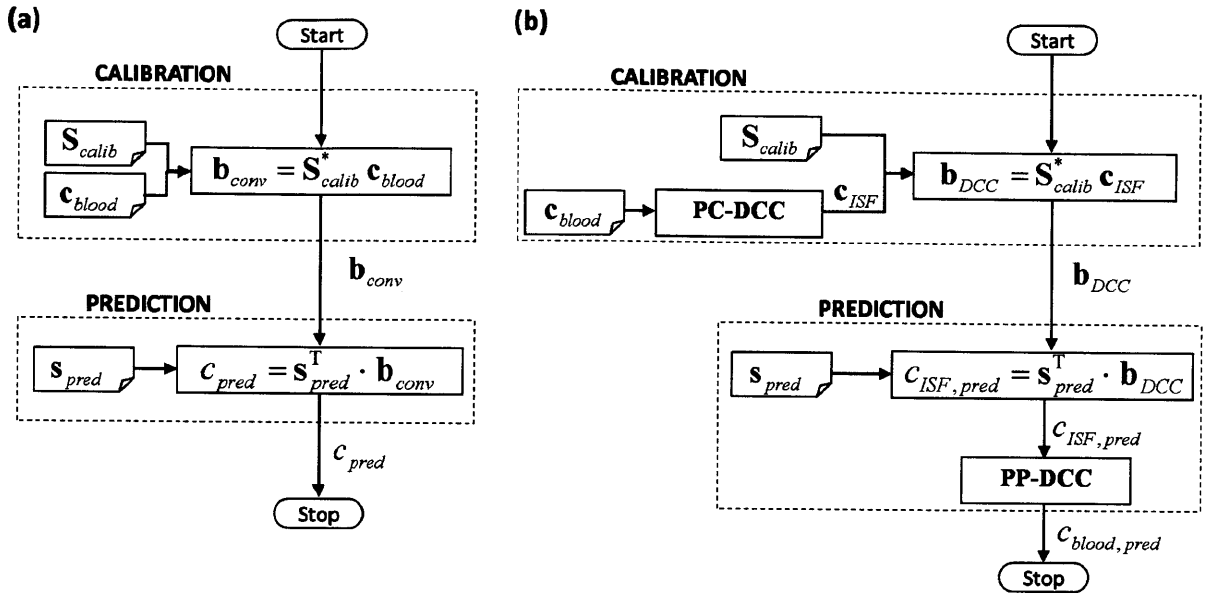


Fig. 7.1 Flowcharts of (a) the conventional implicit and (b) DCC calibration methods. $\mathbf{S}_{calib}$, $\mathbf{c}_{blood}$, and $\mathbf{s}_{pred}$ represent the calibration spectra, reference blood glucose concentrations in the calibration samples, and the spectrum acquired from the prediction sample, respectively. For the conventional calibration method, $\mathbf{b}_{conv}$ and c_{pred} give the regression vector and the predicted concentration, respectively. For DCC calibration, $\mathbf{b}_{DCC}$ represents the developed regression vector. $c_{ISF,pred}$ and $c_{blood,pred}$ are the intermediate ISF glucose estimate and the final blood glucose prediction. PC-DCC is the pre-calibration transformation of blood glucose concentrations into the corresponding ISF glucose values. PP-DCC transforms the predicted ISF glucose concentration into the blood glucose value. Note that the conventional calibration scheme does not differentiate between the blood and ISF glucose concentrations.

For the subsequent analysis, we assume that the sampling volume (i.e. the volume of tissue probed by the NIR light) is a subset of the interstitial fluid space. This assumption is primarily based on the fact that the ISF constitutes nearly 45% of the volume fraction of the human skin in contrast to the blood vessels which contribute about 5% of the skin volume [22]. In Sec.-7.5, we

will revisit this assumption and characterize its impact on the proposed calibration model. With the aid of this simplifying assumption, we can write a simple two-compartment mass transfer model for the ISF volume V_{ISF} [15, 20, 23]:

$$V_{ISF}\frac{dc_{ISF}}{dt} = k_M A(c_{BG} - c_{ISF}) - k_U V_{ISF} c_{ISF}, \tag{7.3}$$

where c_{ISF} and c_{BG} are the ISF and blood glucose concentrations (mol/cm^3) respectively, k_M is the glucose mass transfer coefficient (cm/s), A is the effective mass transfer surface area (cm^2) and k_U is the rate of glucose uptake by the neighboring cells (1/s). The effect of insulin on the uptake term has been ignored. This approximation is consistent with the observed result that the glucose levels in (subcutaneous) ISF are largely unaffected by the local insulin concentration [20]. In fact, the uptake term itself has been observed to be very small for subcutaneous glucose sensing [15]. This is attributed to the fact that sub-cutaneous tissue cells, as opposed to muscle or adipose tissue cells, is unlikely to have significant glucose uptake, even in the presence of high insulin concentration. Consequently, we will ignore the uptake term in further analysis. Eq. (7.3) can then be simplified to:

$$c_{BG} = c_{ISF} + \alpha\frac{dc_{ISF}}{dt}, \tag{7.4}$$

where α ($= \frac{V_{ISF}}{k_M A}$) is a lumped mass-transfer parameter having units of time. The parameter α provides a measure of the physiological lag time arising from the diffusion process and is henceforth called the characteristic lag time constant. This equation provides the ability to construct blood glucose estimates based on the spectroscopy-based ISF glucose prediction values and *a priori* knowledge of the lag time constant in the sample.

A first-order accurate estimate of the blood glucose concentration can be obtained by using a finite difference approximation for the derivative term of Eq. (7.4):

$$c_{BG}(t) = c_{ISF}(t) + \alpha \frac{c_{ISF}(t) - c_{ISF}(t - \Delta t)}{\Delta t} \tag{7.5}$$

where Δt is the time interval at which spectroscopic measurements (and thus, ISF glucose concentration estimates) are obtained. The above equation gives the discrete transformation equation of the post-prediction (PP-DCC) step, which can be applied in real-time. There are a couple of points worth noting about the application of Eq. (7.4). Firstly, it is evident from the above equation that at least two spectroscopic predictions of ISF glucose (at t and $t-\Delta t$) are necessary in order to determine the blood glucose concentration at time t. In practice, it is beneficial to perform multiple spectroscopic acquisitions so that the corresponding blood glucose estimates can be averaged to ensure less fluctuation in the predicted blood glucose value. Secondly, the time interval Δt, at which the spectroscopy based ISF glucose predictions should be performed, needs to be ascertained. Although spectroscopic acquisitions can be performed rapidly, having a very small Δt is not too useful as it may fail to capture the changes in the glucose levels due to the slow diffusion kinetics – thereby rendering the DCC approach ineffective. On the other hand, too large a time interval would introduce substantial errors in the derivative term. Using these limiting cases as guidelines, one might optimize the value of Δt with a starting point given by a fraction of the typical physiological lag time (~5-10 min).

The other important portion of the proposed scheme is the pre-calibration (PC-DCC) step. In order to obtain consistency in the calibration model, we need to convert the measured blood glucose concentrations to the corresponding ISF glucose values. To perform this transformation, we write Eq. (7.3) in its integral form:

$$c_{ISF}(t_f) = c_{ISF}(t_i)\exp\left(-\frac{1}{\alpha}(t_f - t_i)\right) + \frac{1}{\alpha}\int_{t_i}^{t_f} c_{BG}(t)\exp\left(-\frac{1}{\alpha}(t_f - t)\right)dt \quad , \tag{7.6}$$

where the definite integral is evaluated from time t_i to t_f. After employing integration by parts for the second term of Eq. (7.6), numerical integration was performed by using Simpson's rule to get the following equation:

$$c_{ISF}(t_f) = c_{BG}(t_f) + \left(c_{ISF}(t_i) - c_{BG}(t_i)\right)\exp\left(-\frac{t_f - t_i}{\alpha}\right) - \frac{t_f - t_i}{6}A \quad , \tag{7.7a}$$

where

$$A = \dot{c}_{BG}(t_i)\exp\left(-\frac{t_f - t_i}{\alpha}\right) + 4\dot{c}_{BG}\left(\frac{t_i + t_f}{2}\right)\exp\left(-\frac{t_f - t_i}{2\alpha}\right) + \dot{c}_{BG}(t_f) \tag{7.7b}$$

Here, $\dot{c}_{BG}(t)$ refers to the derivative of c_{BG} evaluated at t. This equation can be readily evaluated by approximating $\dot{c}_{BG}(t)$ via a first-order finite difference approximation similar to that employed in Eq. (7.5) for the PP-DCC step. The resultant discretized version of Eq. (7.7) gives the necessary transformation equation for the PC-DCC step. The primary challenge in the evaluation of this equation lies in having *a priori* knowledge of the initial ISF glucose value, i.e. at the start of the time window [t_i, t_f]. Generally, such information is not available for an arbitrary time window. However, during a spectroscopic calibration study such as glucose clamp or tolerance test, one can ensure that the ISF glucose and blood glucose values are completely in equilibrium at the start of the experiment by restricting the glucose intake of the subject prior to the measurements. For example, a typical oral glucose tolerance test (OGTT) protocol stipulates that the patient must fast for 8-14 hours before the study [24]. This initial condition enables successful evaluation of Eq. (7.7) for the first time window. For each subsequent evaluation, the

ISF glucose value at time t_f for the previous window is inputted as the initial value (at time t_i) for the current window. Evidently, the shorter the time window over which the evaluation is performed, the higher the accuracy of the determined ISF glucose concentrations. Nevertheless, the time window cannot be shortened below a lower bound, governed by the maximum permissible frequency of blood withdrawal from a human subject. Most research laboratories, for example, sample blood glucose at time intervals ranging from 2.5-10 minutes [1, 25]. To determine the values of the concentrations and their derivatives at intermediate points, spline interpolation can be employed.

7.3 Prediction uncertainty arising from physiological lag

In order to quantify the precision of spectroscopy based calibration models, Lorber and Kowalski derived an elegant prediction error formula, which describes the error propagation for linear multivariate prediction algorithms [17]. Our laboratory has previously derived analytical expressions for uncertainty in concentration prediction for the specific case where noise in the prediction dataset (spectra) is the dominant source of error [18, 19]. This case is important, as in most biomedical applications constraints on the acquisition time in prediction samples cause the noise in the prediction dataset to be significantly higher than that observed in the calibration dataset (where acquisition times can typically afford to be much longer). In such cases, it can be assumed that an accurate calibration model can be achieved by developing it on calibration samples in which sufficiently high signal-to-noise ratio (SNR) can be attained. Under such conditions, the limiting uncertainty in the predicted concentrations arises from the spectral overlap between the analyte of interest and the other tissue constituents, and the measurement noise in the spectra acquired from the prediction samples. Mathematically, the spectroscopic uncertainty for the analyte of interest (glucose) is given by Δc_s [19]:

$$\hat{c} = (\mathbf{s} \pm \Delta\mathbf{s})^{T} \cdot \mathbf{b} = \mathbf{s}^{T} \cdot \mathbf{b} \pm \Delta\mathbf{s}^{T} \cdot \mathbf{b} = c \pm \Delta c_s \tag{7.8a}$$

where

$$\Delta c_s = \frac{\sigma}{s_g} olf \tag{7.8b}$$

Here, $\hat{c}$ and c are the estimated and actual analyte concentrations in the prediction sample, respectively and $\Delta\mathbf{s}$ represents the spectral noise in the prediction spectrum **s**. As the modeling noise is ignored in computation of the limiting uncertainty, **b** represents the ideal (noise-free) regression vector for glucose. σ is a measure of the noise magnitude in the prediction spectrum, s_g quantifies the signal strength of glucose at unit concentration, and *olf* indicates the amount of overlap between glucose and the other spectral interferents (such as proteins, lipids, and water).

In addition to the spectroscopic uncertainty, there exists a prediction uncertainty for transcutaneous glucose measurements that arises from the physiological lag between blood and ISF glucose levels. Even if the calibration models are developed under conditions in which the blood and ISF glucose concentrations are in equilibrium (such as those obtained by employing glucose/insulin clamps), the predicted concentrations will still contain uncertainties due to the unaccounted physiological lag in the prediction samples. We present an error propagation analysis to determine the limiting uncertainty in concentration prediction due to the physiological lag, with and without DCC. In the spirit of deriving a simple analytical expression, we will assume that the developed calibration model itself is accurate, i.e. devoid of noise and lag-related errors (similar to the underlying assumptions in [18, 19]).

7.3.1 Limiting uncertainty for conventional calibration

When the modeling noise is ignored, Eq. (7.8) provides the relationship between the estimated and the actual glucose concentrations. However, for *in vivo* prediction, there will be a lag between the instantaneous blood and ISF glucose values in the sample, where the latter is measured by the prediction spectrum. Taking this into account, we can re-write Eq. (7.8) in terms of the estimated ($\hat{c}_{ISF}$) and actual (c_{ISF}) ISF glucose concentrations:

$$\hat{c}_{ISF} = (\mathbf{s} \pm \Delta\mathbf{s})^{\mathrm{T}} \cdot \mathbf{b} = c_{ISF} \pm \Delta c_s . \tag{7.9}$$

Based on the two-compartment model, the actual blood glucose concentration, c_{BG}, can be determined from the actual ISF glucose concentration using Eq. (7.4), given the correct lag time constant for the prediction sample (α_{actual}). Substituting the value of the actual ISF glucose concentration from Eq. (7.9) into Eq. (7.4), we obtain:

$$c_{BG} = c_{ISF} + \alpha_{actual}\frac{dc_{ISF}}{dt} = \hat{c}_{ISF} + \alpha_{actual}\frac{d\hat{c}_{ISF}}{dt} \pm \Delta c_s \tag{7.10}$$

However, the conventional models report the estimated ISF glucose concentration as the blood glucose concentration ($\hat{c}_{BG}$) in the prediction sample:

$$\hat{c}_{BG} = \hat{c}_{ISF} . \tag{7.11}$$

Substituting Eq. (7.11) into Eq. (7.10) and re-arranging, we obtain:

$$\hat{c}_{BG} = c_{BG} - \alpha_{actual}\frac{dc_{ISF}}{dt} \pm \Delta c_s \tag{7.12}$$

Equation (7.12) implies that under the conventional calibration framework, the limiting uncertainty in the concentration estimate has two separate contributions: (i) the uncertainty

resulting from the measurement noise in the prediction spectrum and the spectral overlap, Δc_s and (ii) the uncertainty due to the glucose physiological lag, $\Delta c_{conv} = \alpha_{actual} \frac{dc_{ISF}}{dt}$. While the former (i) is a well-known quantity, the latter uncertainty (ii) has not been examined before.

Figure 7.2 illustrates the two contributing factors of the prediction uncertainty. In this figure, the simulated blood and ISF glucose data in panel (A) (which mimics the glucose profiles obtained from a tolerance test) are plotted against each other to construct the solid line curve in panel (B). It is evident that the physiological lag between the blood and ISF glucose profiles in (A) introduces a hysteresis-like closed loop behavior when blood glucose is plotted against ISF glucose, showing the lack of a one-to-one correspondence between the glucose concentrations in the two compartments. For example, we observe that given an ISF concentration of 148 mg/dl at point P, the actual blood glucose concentration could be either 132 mg/dl (Q) or 158 mg/dl (T). However, conventional methods that have the underlying assumption of a constant blood-to-ISF glucose gradient would predict 148 mg/dl (R), giving rise to a significant error in prediction. Specifically, when the glucose levels are increasing, the blood glucose concentrations are greater than the corresponding ISF glucose concentrations. This set of values is represented by the points on the concave upwards curve (labeled "RISE"). Similarly, the set of values obtained during the falling phase is represented by the concave downwards curve (labeled "FALL"). The lag uncertainty Δc_{conv} in the predicted blood glucose concentration for the conventional calibration model is given by the distance between points Q and R. The uncertainty due to systematic noise and spectral overlap in the prediction spectrum is marked as Δc_s.

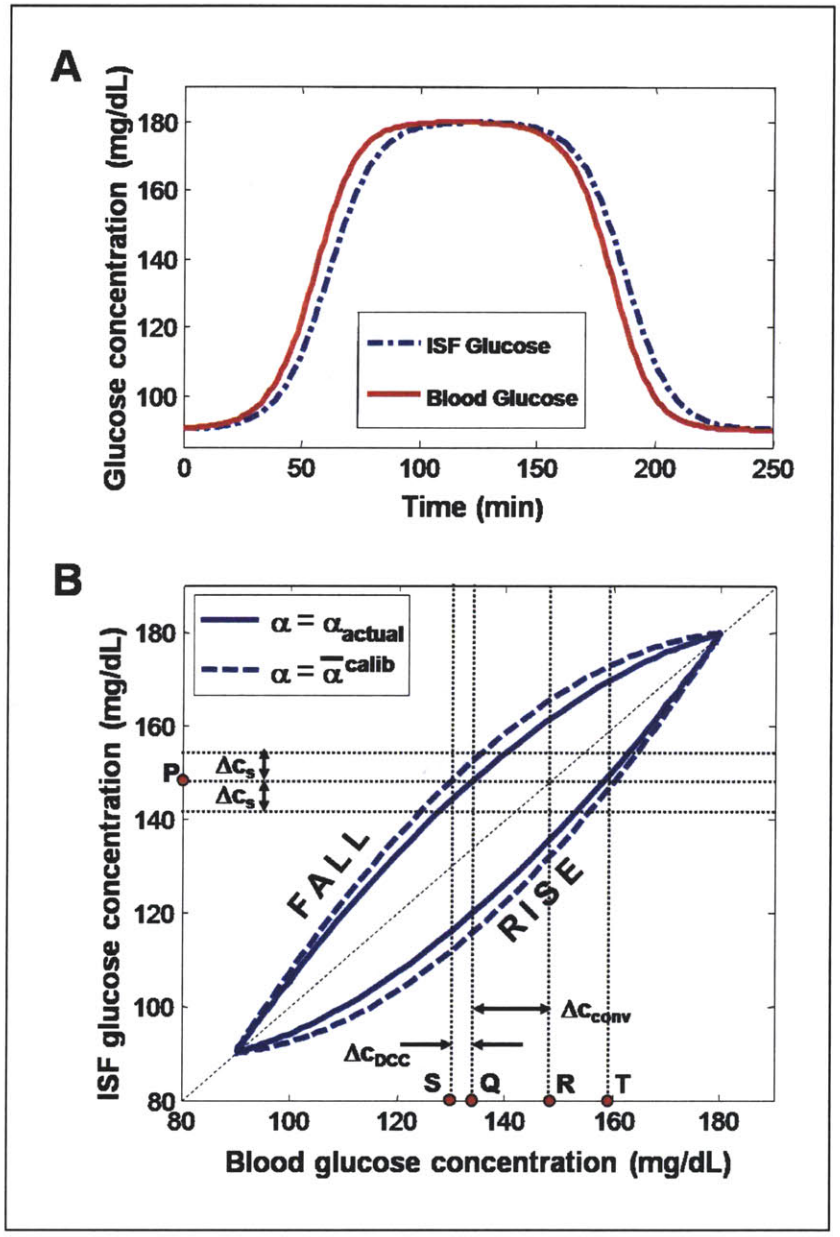


Fig. 7.2 **(A) A schematic representation of blood and ISF glucose concentration profiles, similar to those obtained during a typical tolerance test. (B) Plot of the ISF vs. blood glucose concentrations shown in panel (A). The solid line curve shows the lack of one-to-one correspondence between the actual ISF and blood glucose relationship, while the dotted line curve represents the approximate relationship estimated by the DCC model. Further details are provided in the text.**

7.3.2 Limiting uncertainty for DCC calibration

In contrast to the conventional model, the DCC scheme explicitly accounts for the physiological glucose dynamics. Specifically, the post-prediction equation (PP-DCC) is used to transform the spectroscopy-based ISF glucose estimate ($\hat{c}_{ISF}$) to a corresponding blood glucose value ($\hat{c}_{BG}$), and this step needs to be considered in evaluating the limiting uncertainty. As the correct lag time constant in the prediction sample is unknown in a real clinical setting, some uncertainty due to the physiological lag is introduced via the PP-DCC step. As explained in Sec.-7.3.1, we employ the (ensemble) average of the lag time constants obtained from the calibration samples to approximate the actual lag time constant in the prediction sample. Based on this approximation, the PP-DCC equation can be re-written as:

$$\hat{c}_{BG} = \hat{c}_{ISF} + \overline{\alpha}^{calib} \frac{d\hat{c}_{ISF}}{dt}, \tag{7.13}$$

where $\overline{\alpha}^{calib}$ refers to the average value of α computed from the calibration samples.

Substituting Eq. (7.9) into Eq. (7.13), we obtain:

$$\hat{c}_{BG} = c_{ISF} \pm \Delta c_s + \overline{\alpha}^{calib} \frac{dc_{ISF}}{dt}. \tag{7.14}$$

The deviation of $\overline{\alpha}^{calib}$ from the actual lag time constant in the prediction sample, α_{actual} can be written as:

$$\overline{\alpha}^{calib} = \alpha_{actual} \pm \Delta\alpha, \tag{7.15}$$

where $\Delta\alpha$ is the error (uncertainty) in the estimation of the lag time constant.

Substituting Eq. (7.15) into Eq. (7.14) and re-writing the first term of the above equation as c_{BG}, we get:

$$\hat{c}_{BG} = c_{BG} \pm \Delta\alpha \frac{dc_{ISF}}{dt} \pm \Delta c_s \tag{7.16}$$

Equation (7.16), which is analogous to Eq. (7.12) for the conventional calibration model, implies that even with DCC, the net uncertainty is a combination of the uncertainties arising from the spectral noise and overlap (Δc_s) and the physiological lag ($\Delta c_{DCC} = \Delta\alpha \frac{dc_{ISF}}{dt}$). However, the primary difference between the two cases – with and without DCC– lies in the magnitude of uncertainty introduced due to the physiological lag. The lag uncertainty for the conventional calibration model case (which is proportional to α_{actual}) is significantly larger than that observed for DCC calibration (which is proportional to $\Delta\alpha$). This can be visualized in Fig. 2(B). The dashed line curve of Fig. 2 (B) connects the points whose co-ordinates are given by the model *estimated* blood and ISF glucose concentrations (in contrast to the solid line curve represents the points whose co-ordinates are given by the *actual* blood and ISF glucose concentrations). Since the exact lag time constant of the prediction sample is unknown, the estimated blood glucose concentrations will differ from the actual blood glucose concentrations by the product of the rate of change in glucose concentration and the estimation uncertainty of the lag time ($\Delta c_{DCC} = \Delta\alpha \frac{dc_{ISF}}{dt}$). It is worth noting that the dashed (DCC estimated) curve is computed using Eq. (7.13), whereas the blood and ISF glucose concentrations of the solid curve are related by Eq. (7.4). From the figure, it is evident that Δc_{DCC}, the distance between points Q and S, is substantially smaller than Δc_{conv}, the distance between points Q and R, as long as the lag time constant used in the DCC model provides a reasonably close approximation to the actual

lag time constant. A quantitative comparison of the two lag uncertainties and the spectroscopic uncertainty is performed in Sec.-7.5.2.

7.4 Materials and Methods

We performed numerical simulations and experimental studies to: (1) demonstrate the improvement in prospective prediction performance of the calibration model on application of DCC and (2) estimate the distribution of the lag time constant in a human population and characterize the prediction uncertainty introduced due to the physiological lag. To accomplish (1), a numerical simulation study was undertaken (Sec.-7.4.1). In this study, ISF and blood glucose concentration datasets, described by Steil *et al.*[20], were used to generate tissue Raman spectra for calibration and prediction. The simulations were also used to understand the relationship between the SNR in the spectral dataset and performance of the conventional and DCC calibration models. In order to investigate the lag time distribution in a human population (2), datasets obtained from our laboratory's clinical studies on human volunteers were employed. Additionally, the human volunteer study was used to determine the limiting uncertainty arising from the physiological glucose dynamics, as described in Sec.-7.4.2.

7.4.1 Numerical Simulations

The dataset used in our numerical simulations was based on blood and ISF glucose concentrations originally measured by Steil *et al.* [20]. In their studies, Steil and co-workers monitored blood and ISF glucose concentrations in non-diabetic human subjects during glucose clamping. After 10–12 hours of overnight fasting, glucose was sequentially clamped at approximately 5, 4.2 and 3.1 mM (1 mM of glucose $\approx$ 18 mg/dL) for 90 minutes each by insulin and glucose infusion, and subsequently allowed to return to euglycemic levels. ISF glucose was

measured by two MiniMed (Medtronic, Inc.) subcutaneous amperometric glucose sensors. Blood was withdrawn at regular intervals for blood glucose measurements using a clinical glucose analyzer. Our analysis uses the blood and ISF glucose concentrations from 90 to 380 minutes after initial insulin and glucose infusion, as shown in Fig. 7.3. The simulated spectra and corresponding blood glucose concentrations is divided into calibration (dataset spanning from 90 to 220 min) and prediction (dataset from 230 to 380 min) sets, respectively.

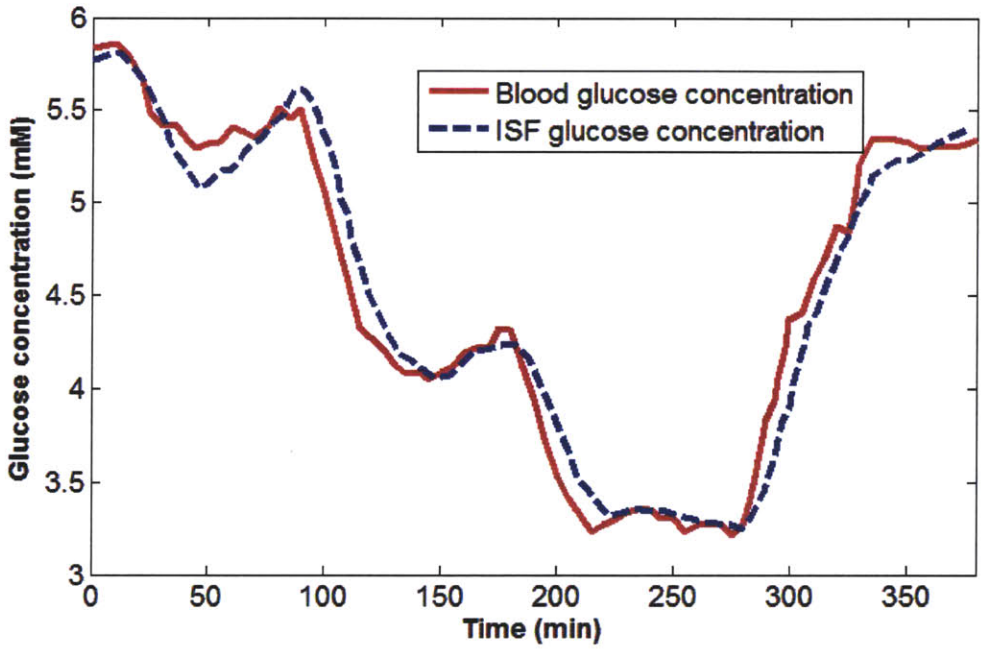


Fig. 7.3 Blood and ISF glucose concentration time profiles measured from a normal human volunteer during insulin-induced hypoglycemia [20]. Glucose was clamped at 5, 4.2 and 3.1 mM and subsequently allowed to return to normoglycemic levels. It is observed that the ISF glucose, measured by subcutaneous amperometric sensors, consistently lags blood glucose concentrations during both rising and falling phases. In contrast, they have nearly identical values during the clamping phases. (Reprinted from Reference 20, Copyright 2005, with permission from Springer Science+Business Media: Diabetologia.)

In our study, simulated Raman spectra are generated by forming weighted linear combinations of the constituent Raman spectra of glucose, creatinine, and urea (as measured by our laboratory Raman system). The weights assigned for glucose (the analyte of interest) are determined by the experimentally measured ISF glucose concentrations of the Steil dataset. The other two constituents (spectral interferents) are assigned weights that randomly varied within 2% of a constant value, in order to mimic the small changes observed in these constituents during typical glucose tolerance and clamping tests. To simulate normal experimental conditions, zero-mean Gaussian white noise (i.e. homoscedastic noise contributions) is added to the mixture spectra at varying levels of signal to noise ratio (SNR of 20-40 dB) to study its effect on prediction performance of the calibration models.

In contrast to the conventional PLS calibration strategy, where the number of loading vectors are optimized, in the DCC calibration the number of loading vectors as well as the lag time constant, α, need to be optimized. To accomplish this, we initially assign default values to α (0 minutes) and the number of loading vectors (2) employed, respectively. Prior to constructing the leave-one-out calibration model, all but one of the reference blood glucose concentrations are converted to the corresponding ISF glucose values using PC-DCC (Eq. (7.7)). This allows the creation of a calibration model based purely on ISF glucose. The developed calibration model, in conjunction with the spectrum of the excluded data point (which constitutes the validation data), is then used to predict the ISF glucose concentration at that point. Subsequently, PP-DCC (Eq. (7.5)) is used to re-transform the predicted ISF glucose concentration to the blood glucose value. This process is repeated till each data point is used once as the validation data. The resultant blood glucose estimates are compared with the actual blood glucose values to give the root-mean-squared error of cross-validation (RMSECV). The whole procedure is iterated for

appropriate ranges of α (0 to 20 min) and number of loading vectors (LV) (2 to 10) to determine the optimal combination of parameters (α_{opt}, LV_{opt}) that yields the minimum RMSECV. This combination of parameters is then used to obtain the PLS regression vector, $\mathbf{b}_{opt}$. Prospective prediction on a separate portion of the data set was performed by taking the scalar product of the prediction spectra with $\mathbf{b}_{opt}$ (Eq. (7.2)). The ISF glucose predictions are re-converted to the blood glucose values using PP-DCC, where α_{opt} is used in place of α in Eq. (7.5). The root-mean-squared error of prediction (RMSEP) is computed from the predicted blood glucose concentrations and the reference blood glucose values.

Conventional PLS calibration and prediction is also performed on the same dataset to compare the relative performance with the DCC model. Twenty simulations are carried out for each value of SNR in the spectral dataset (both calibration and prediction) to establish the mean and standard deviation of the prospective prediction errors.

7.4.2 Human subject studies

To investigate the lag time distribution in a human population, clinical datasets consisting of blood glucose concentrations and tissue Raman spectra are used. The acquisition of the clinical data was originally described in one of our laboratory's previous publications [26] and has also been detailed in Chapter-3. Datasets from human subjects exhibiting motional artifacts, inadequate SNR in the acquired spectra, and impaired glucose tolerance characteristics are excluded from further analysis. For the selected volunteer datasets, DCC calibration is performed using a leave-one-out cross-validation routine on the measured Raman spectra and reference blood glucose concentrations to determine the optimal value of α for each individual. In addition, conventional PLS calibration is also performed on the same datasets to compare the

resultant cross-validation errors. The cross-validation procedures in both cases remain the same as that described in Sec.-7.4.1, except that experimentally measured Raman spectra are used in place of the simulated Raman spectra. The mean and standard deviation of α determined from the human subjects are used to approximate α_{actual} and $\Delta\alpha$ for the quantification of uncertainty due to physiological lag for the DCC and conventional calibration schemes respectively (Sec.-7.3). These uncertainty estimates are also compared with the spectroscopic uncertainty, Δc_s.

7.5 Results and Discussion

7.5.1 Numerical simulations

Numerical simulations are used to compare the prospective prediction capability of the conventional and DCC calibration models. We observe that DCC implementation is found to reduce the cross-validation error, RMSECV, of the simulated dataset from 0.15 mM to 0.04 mM, when the measured ISF glucose concentrations are used for computing the cross validation errors. These simulation results are obtained for a SNR of 40 dB. This result is represented in Fig. 7.4, where the measured ISF glucose concentrations are plotted together with cross-validated glucose concentrations from the conventional and DCC calibration models. It is evident that the ISF glucose concentration profile generated with DCC closely matches the measured ISF glucose concentrations, while that generated without DCC shows significantly larger deviations. The cross-validation routine also optimized the lag time constant for the DCC calibration model and the number of loading vectors for both models. For this dataset, the characteristic lag time constant α_{opt} is determined to be 6.1 minutes, in agreement with the experimentally observed values of 6–8 minutes [20].

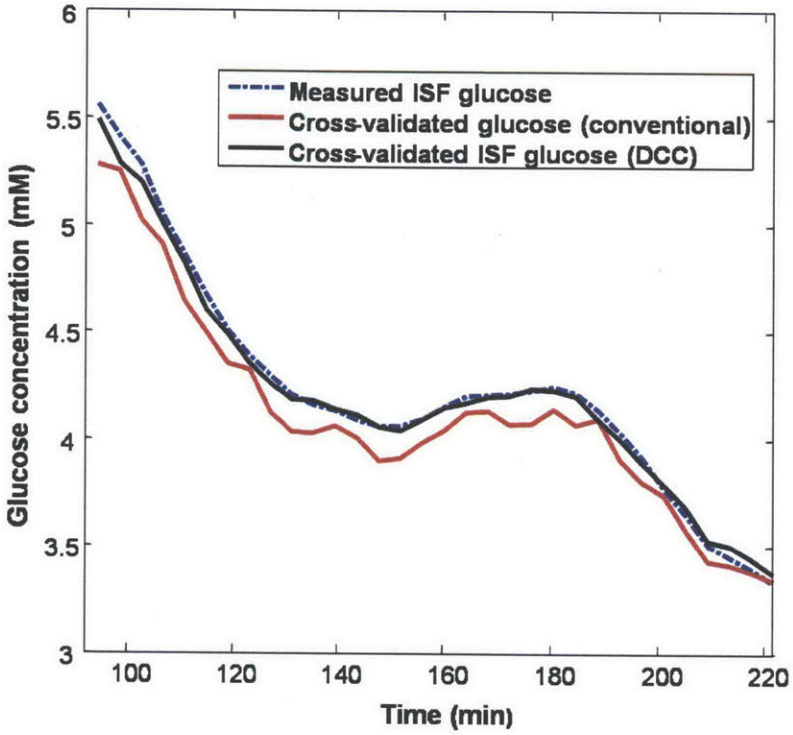


Fig. 7.4 Cross-validation results of conventional (red) and DCC (black) calibration methods applied on the simulated dataset. The measured ISF glucose concentration values are given by the blue dotted line. In the DCC calibration process, the lag time constant α_{opt} was optimized to be 6.1 minutes.

When the calibration models are applied prospectively to the prediction dataset (i.e. the rising phases of the glucose clamps), the DCC model (RMSEP = 0.14 mM) is found to exhibit significantly improved prediction accuracy compared with the conventional PLS scheme (RMSEP = 0.28 mM). Figure 7.5 below shows the results of prospective prediction, in which the measured blood glucose concentration profile is plotted alongside of the prediction profiles, with and without DCC. This demonstrates how calibration during non-equilibrium conditions leads to systematic errors giving rise to much higher prediction errors (Fig. 7.5) than estimated during cross-validation (Fig. 7.4). In fact, in the presence of such errors, the predicted glucose

concentration may retain no statistically significant correlation with the true glucose concentrations during rapidly rising and declining glucose concentrations.

Potentially, one could achieve an even closer correlation with the measured blood glucose concentration profile by smoothing out the noisy fluctuations observed in the concentration profile of the DCC prediction of Fig. 7.5. However, such smoothing algorithms were not employed as they might introduce artifacts and additional delays to the concentration profile that are unrelated to glucose equilibration [20].

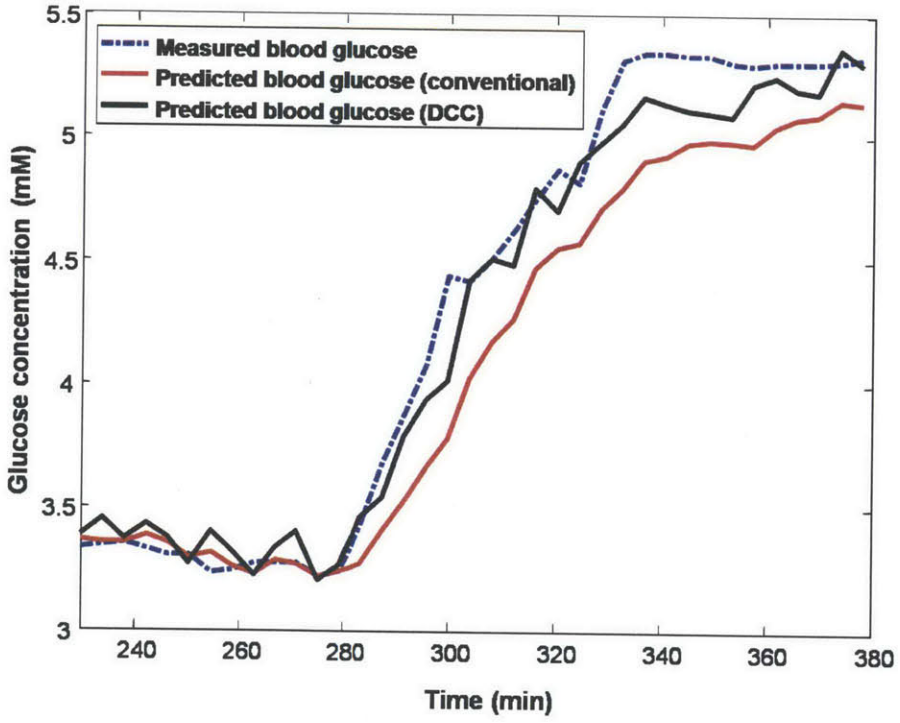


Fig. 7.5 Prospective prediction results of conventional (red) and DCC-based (black) calibration methods applied on the simulated dataset. The measured blood glucose concentration values are given by the blue dotted line.

The accuracy of blood glucose concentration prediction with and without DCC was also compared at varying levels of SNR in the spectral dataset. Figure 7.6 shows a plot of the RMSEP

of blood glucose prediction with conventional and DCC calibration models as a function of SNR. In both cases, increase in noise level corresponded to an increase in error values, as expected. However, under all tested values of SNR, the prediction error of DCC calibration models was consistently smaller compared to that of the conventional PLS models.

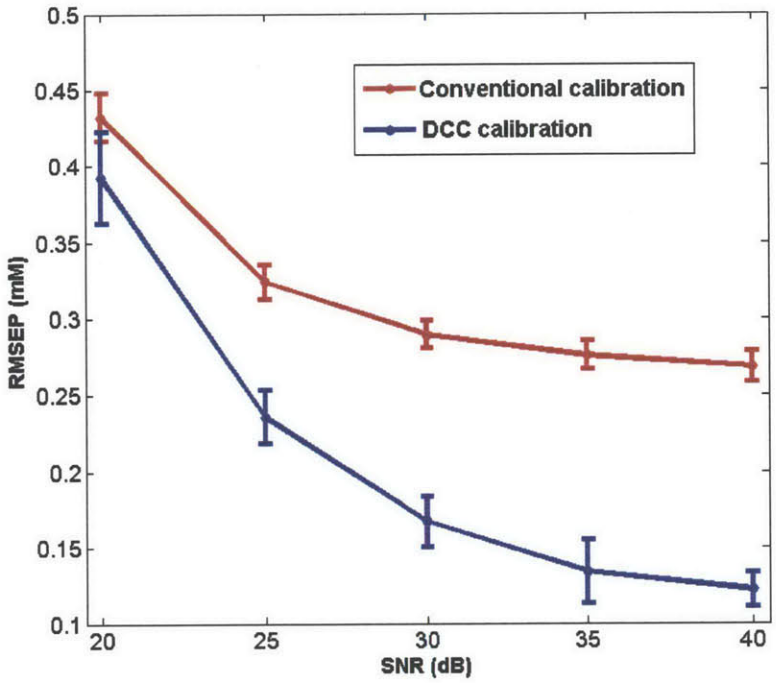


Fig. 7.6 Plot of RMSEP obtained for conventional (red) and DCC (blue) calibration models, applied on the simulated dataset, as a function of increasing SNR. The error bars represent the standard deviation of RMSEP for 20 iterations.

Evidently for DCC implementation, a time series of glucose measurements (such as those obtained from glucose tolerance tests) are required from human subjects included in the calibration study for characterization of the glucose kinetics. It is not possible to develop a consistent calibration model based only on single spectroscopic measurements from multiple human subjects; as such measurements do not provide information on the glucose kinetics.

It is also worth noting that application of an enhanced calibration scheme, such as support vector machines [27, Chapter-5] or hybrid calibration methods [28, 29], alone would not alleviate inconsistencies in the calibration models as they do not address the lack of one-to-one correspondence between the ISF and blood glucose concentrations (Fig. 7.2). Nevertheless, when used in conjunction with DCC, such calibration schemes may potentially further improve the prediction accuracy.

7.5.2 Human subject studies

Table 1 lists the results of the leave-one-out cross-validation on the dataset from each human subject, using DCC calibration as well as the conventional PLS routine. We observe that, on average, the RMSECV for the blood glucose concentrations of the human subjects reduces on application of DCC by 15.5%, with a maximum reduction of 28.6%. This demonstrates the applicability of DCC in clinical situations, where the number of tissue components probed is vastly greater than the three constituents (glucose, creatinine and urea) employed in our simulations.

Table 7.1 Summary of cross-validation results of conventional and DCC calibration models applied on the human volunteer datasets (Sec.-7.4.2).

Volunteer	No. of data points	DCC Model		Conventional Model	% change in RMSECV
		α_opt (min)	RMSECV (mM)	RMSECV (mM)	
1	25	9.5	0.87	1.07	19.10
2	26	11.5	0.52	0.72	27.56
3	26	10.5	0.70	0.97	28.62
4	30	11.1	1.02	1.08	5.74
5	25	8.4	0.68	0.79	13.35
6	26	8.4	0.64	0.82	22.00
7	25	7.5	0.42	0.51	17.61
8	29	8.1	0.76	0.87	12.44
9	31	8.3	0.96	1.00	3.41
10	27	12	0.50	0.53	5.25

As mentioned in Sec.-7.2, it has been assumed in formulating the DCC scheme that the sampling volume is a subset of the interstitial fluid space. This assumption was based on: (i) NIR light has a penetration depth of ~1 mm in skin tissue and (ii) the blood vessels contribute only 5% to the total skin volume [22] with the outermost epidermis being completely avascular. However, a small fraction of the inelastically scattered (Raman) light arises from the glucose residing in the blood compartment. This results in an effective reduction in the value of the lag time constant α, as determined by our DCC model. Nevertheless, our results demonstrate that DCC successfully models the clinical human volunteer data, even though a small Raman contribution from the blood glucose component is present. This shows that the DCC approach is effective in improving consistency in the calibration model and thus in prospective prediction, as long as the spectral contribution of blood glucose is small compared to that of ISF glucose.

Although the mass transfer formulation would remain valid, a weighted combination of the two glucose values may need to be used to correlate with the spectral measurements, if the spectral contribution from blood glucose is comparable to that from the ISF glucose.

The results also seem to suggest that the value of the lag time is fairly constant for the tested human volunteer population. This is established by the lag time distribution obtained from the clinical data, where the mean of α, 9.5 min, is significantly larger than the standard deviation, 1.6 min. The relative constancy of α indicates that the mean lag time of the calibration set provides a fairly accurate approximation to the lag time of any prospective subject, on whom the algorithm has not been applied before. Of course, this is a fairly small population set from which it is not possible draw much meaningful statistics and our current large-scale clinical studies (Chapter 7.6) should provide a better understanding of the relative constancy - or lack thereof - of the lag time constant.

To the best of our knowledge, this is the first time that the lag time distribution in a human population has been measured by optical techniques. Previous attempts with subcutaneous amperometric sensors have been observed to have significant sensor specific lag, which obscures the precision of the physiological lag measurements [1]. In addition to the sensor specific lag of the sub-cutaneous amperometric monitors, the difference in lag time constants observed from the numerical simulations using the Steil dataset (6.1 min) and our clinical studies on human volunteers (9.5±1.6 min) can be attributed to: (a) the difference in the composition of glucose (i.e. the proportion of blood and ISF) sampled by the spectroscopic and amperometric sensors and (b) the variations in the population demographics studied in the two cases.

Although most current research including our own study on human subjects reported above indicates a reasonably constant value of α for a single individual, some research groups have

previously suggested that the response times between blood and ISF glucose may be different for the rising and falling phases. These investigators have claimed that ISF glucose falls in advance of blood glucose during the time of declining glucose levels [30-32]. In such situations, a modified DCC model can be implemented by employing two distinct α's during the rising and falling phases through a piece-wise application of Eq. (7.5) and (7.7).

The mean and standard deviation of α obtained on the human volunteer dataset were also employed in determining the physiological lag uncertainties for the DCC and conventional calibration schemes. For the conventional schemes, the uncertainty is calculated as $\alpha_{actual}\frac{dc_{ISF}}{dt}$, where the mean α of the human volunteers is used for α_{actual}. At times of reasonably rapid increase in glucose levels, the concentration of glucose in either compartment may change by ~2 mg/dL/min (0.11 mM/min) [5]. Plugging in these values, we find that for conventional calibration, the prediction uncertainty due to lag amounts to approximately 1.06 mM. For the DCC model, the uncertainty due to lag can be computed by $\Delta\alpha\frac{dc_{ISF}}{dt}$, where $\Delta\alpha$ is approximated by the standard deviation of α obtained from the human volunteer dataset. Based on this value of $\Delta\alpha$, we obtain a lag uncertainty of 0.18 mM for the DCC calibration method. The calculated values project to an approximately six-fold reduction in the lag uncertainty on application of DCC.

Previously, we had estimated using tissue phantom studies that the spectroscopic uncertainty for glucose using our Raman instrument was 1.04 mM (obtained for σ=61.03 (photon counts), s_g=83.74 (photon counts/mM) and olf_g=1.43 in Eq. (7.8b) [19]). Prior to this work, this spectroscopic uncertainty was considered to be the limit of detection, i.e. the smallest

concentration at which glucose could be detected in the tissue. For the specific case of non-invasive glucose detection this does not provide the full picture, as it ignores the lag uncertainty. For example, when conventional calibration schemes are employed, our results suggest that the uncertainty due to lag (1.06 mM) is comparable to the spectroscopic uncertainty (1.04 mM), especially at times of rapid changes in glucose levels. Consequently, the resultant limit of detection, which is sum of the uncertainties arising from spectroscopic considerations and physiological lag (2.1 mM), may be nearly twice that of the previously accepted value (1.04 mM). On the other hand, the net uncertainty on application of DCC (1.22 mM) predominantly arises from spectroscopic considerations (SNR and spectral overlap) and the glucose kinetics plays only a minor role.

7.6 Summary

The presence of a physiological lag between glucose in the blood and ISF compartments must be considered in developing an accurate spectroscopy-based calibration model for predicting blood glucose concentrations. We have presented a mass transfer model based correction scheme that explicitly accounts for the glucose kinetics. The proposed dynamic concentration correction (DCC) enables us to employ the reference concentrations that are appropriate for the acquired spectra in developing the calibration model – a key step which has not been previously considered. In particular, the resulting improvement in blood glucose estimates should enhance the spectroscopic ability to correctly determine hypoglycemia and even predict impending hypoglycemia based on the rate of change in glucose concentration. Furthermore, we have demonstrated that the prediction uncertainty due to physiological lag, which is comparable in magnitude to the uncertainty arising from noise and non-orthogonality in the spectral dataset, can be reduced substantially by employing DCC.

Our laboratory is currently performing a clinical study to characterize the glucose kinetics in a larger population of human subjects, both with and without diabetes. It is well-known that inadequate glycemic control causes microvascular and macrovascular changes [33], which may in turn affect the physiological lag time. We expect that this study will provide further details about such changes and the ability of spectroscopy to diagnose similar diabetes-related complications. In addition, the clinical study across a larger and more diverse population would lead to a better understanding of the applicability of the DCC model as well as the relative constancy of the lag time across different population segments. This clinical study is also expected to shed light on tissue site selection for spectroscopic sensing, based on a combination of skin heterogeneity and glucose kinetics parameters. In addition, our future research will focus on incorporating control theory elements into the current DCC model to transform this present formulation into one which can dynamically fine-tune its parameters on the fly based on successive spectral measurements.

Acknowledgements

This work was performed in conjunction with Chae-Ryon Kong. Professor Steve Jacques' comments on focusing on the important physiological problems are duly acknowledged. Professor Mark Bathe's desire to see the error bars on the graphs led to the implementation of the SNR-based analysis of the DCC and conventional models. Many of the figures and text were reproduced from the journal article [34] with permission from American Chemical Society (ACS).

7.7 References

(1) Boyne, M. S.; Silver, D. M.; Kaplan, J.; Saudek, C. D. *Diabetes* **2003**, 52, 2790-2794.

(2) Kulcu, E.; Tamada, J. A.; Reach, G.; Potts, R. O.; Lesho, M. J. *Diabetes Care* **2003**, 26, 2405-2409.

(3) Rebrin, K.; Steil, G. M. *Diabetes Technology and Therapeutics* **2000**, 2, 461-472.

(4) Thennadil, S. N.; Rennert, J. L.; Wenzel, B. J.; Hazen, K. H.; Ruchti, T. L.; Block, M. B. *Diabetes Technology and Therapeutics* **2001**, 3, 357-365.

(5) Cengiz, E.; Tamborlane, W. V. *Diabetes Technology and Therapeutics* **2009**, 11, S-11-16.

(6) Zierler, K. *American Journal of Physiology* **1999**, 276, E409–E426.

(7) Koschinsky, T.; Heinemann, L. *Diabetes/Metabolism Research and Reviews* **2001**, 17, 113–123.

(8) McGrath, J. A.; Eady, R. A.; Pope, F. M. *Rook's Textbook of Dermatology*, Blackwell Publishing, **2004**.

(9) Bolinder, J.; Ungerstedt, U.; Arner, P. *Diabetologia* **1992**, 35, 1177-1180.

(10) Reach, G.; Wilson, G. S. *Analytical Chemistry* **1992**, 64, 381A-386A.

(11) Lonnroth, P.; Jansson, P. A.; Smith, U. *American Journal of Physiology* **1987**, 253, E228–E231.

(12) The Diabetes Control and Complications Trial Research Group New England Journal of Medicine 1993, 329, 977-986.

(13) The Diabetes Control and Complications Trial Research Group Diabetes Care 1995, 18, 1415-1427.

(14) Heise, H. M.; Marbach, R.; Koschinsky, T. H.; Gries, F. A. *Artificial Organs* **1994**, 18, 439-447.

(15) Schmidtke, D. W.; Freeland, A. C.; Heller, A.; Bonnecaze, R. T. *Proceedings of National Academy of Sciences* **1998**, 95, 294-299.

(16) Freeland, A. C.; Bonnecaze, R. T. *Annals of Biomedical Engineering* **1999**, 27, 525–537.

(17) Lorber, A.; Kowalski, B. *Journal of Chemometrics* **1988**, 2, 93-109.

(18) Berger, A. J.; Feld, M. S. *Applied Spectroscopy* **1997**, 51, 725-732.

(19) Ščepanovic, O. R.; Bechtel, K. L.; Haka, A. S.; Shih, W. C.; Koo, T. W.; Berger, A. J.; Feld, M. S. *Journal of Biomedical Optics* **2007**, 12, 064012.

(20) Steil, G. M.; Rebrin, K.; Hariri, F.; Jinagonda, S.; Tadros, S.; Darwin, C.; Saad, M. F. *Diabetologia* **2005**, 48, 1833–1840.

(21) Brereton, R. G. *Applied Chemometrics for Scientists*; John Wiley & Sons Ltd,: Chichester, West Sussex, England, **2007**.

(22) Roe, J. N.; Smoller, B. R. *Critical Reviews in Therapeutic Drug Carrier Systems* **1998**, 15, 199-241.

(23) Rebrin, K.; Steil, G. M.; van Antwerp, W. P.; Mastrototaro, J. J. *American Journal of Physiology* **1999**, 277, E561–E571.

(24) American Diabetes Association, *Diabetes Care* **2009**, 32, S13-S61.

(25) Maruo, K.; Oota, T.; Tsurugi, M.; Nakagawa, T.; Arimoto, H.; Tamura, M.; Ozaki, Y.; Yamada, Y. *Applied Spectroscopy* **2006**, 60, 441-449.

(26) Enejder, A. M. K.; Scecina, T. G.; Oh, J.; Hunter, M.; Shih, W.-C.; Sasic, S.; Horowitz, G.; Feld, M. S. *Journal of Biomedical Optics* **2005**, 10, 031114.

(27) Thissen, U.; Ustun, B.; Melssen, W. J.; Buydens, L. M. C. *Analytical Chemistry* **2004**, *76*, 3099-3105.

(28) Berger, A. J.; Koo, T. W.; Itzkan, I.; Feld, M. S. *Analytical Chemistry* **1998**, *70*, 623-627.

(29) Shih, W.-C.; Bechtel, K. L.; Feld, M. S. *Analytical Chemistry* **2007**, *79*, 234-239.

(30) Sternberg, F.; Meyerhoff, C.; Mennel, F. J.; Mayer, H.; Bischof, F.; Pfeiffer, E. F. *Diabetologia* **1996**, 39, 609-612.

(31) Aussedat, B.; Thome-Duret, V.; Reach, G.; Lemmonier, F.; Klein, J. C.; Hu, Y.; Wilson, G. S. *Biosensors and Bioelectronics* **1997**, 12, 1061-1071.

(32) Thome-Duret, V.; Reach, G.; Gangnerau, M. N.; Lemonnier, F.; Klein, J. C.; Zhang, Y.; Hu, Y.; Wilson, G. S. *Analytical Chemistry* **1996**, 68, 3822-3826.

(33) Reusch, J. E. B. *Journal of Clinical Investigation* **2003**, 112, 986-988.

(34) Barman, I.; Kong, C. R.; Singh, G. P.; Dasari, R. R.; Feld, M. S. *Analytical Chemistry* **2010**, 82, 6104-6114.

CHAPTER 8

DESIGN OF MINIATURZED BLOOD GLUCOSE MONITOR

While Raman spectroscopy provides a powerful tool for non-invasive and real time diagnostics of biological samples, its translation to the clinical setting has been impeded by the lack of robustness of spectroscopic calibration models and the size and cumbersome nature of conventional laboratory Raman systems. Linear multivariate calibration models employing full spectrum analysis are often misled by spurious correlations, such as system drift and covariations among constituents. In addition, such calibration schemes are prone to overfitting, especially in the presence of external interferences that may create non-linearities in the spectra-concentration relationship.

To address both of these issues, we incorporate residue error plot-based wavelength selection, to eliminate uninformative regions of the spectrum, into non-linear support vector regression (SVR), which enables appropriate modeling of the curved effects. Using glucose detection in tissue phantoms as a representative example, we show that even a substantial reduction in the number of wavelengths analyzed using SVR lead to calibration models of equivalent prediction accuracy as linear full spectrum analysis. With clinical datasets obtained from human subject studies, we also demonstrate the prospective applicability of the selected wavelength subsets without sacrificing prediction accuracy.

The surprising transferability of the selected wavelength subsets, particularly for the support vector regression models, opens new avenues for calibration transfer using a few important spectral features only. This has important implications for calibration maintenance and transfer

perspective as well as for reduction of acquisition time in serial acquisition systems. This approach also enables the development of a miniaturized clinical system that is programmed for excitation and detection of these wavelength subsets. To implement this in a miniaturized clinical device, we propose system designs featuring tunable excitation and detection components. The proposed systems significantly reduce the spatial footprint and the necessary acquisition time. We anticipate that such systems will play a vital role in disease diagnosis as well as in pharmaceutical applications and forensic analysis.

8.1 Background and significance

Noninvasive diagnostic tools are of great importance in biomedicine as they dramatically reduce the cost and inconvenience associated with blood withdrawals and tissue biopsies. By virtue of its high chemical specificity and capability for reagentless detection of the sample constituents, Raman spectroscopy has received considerable interest in the biomedical community for detection of cancerous lesion [1, 2], atherosclerotic plaque [3] and diabetes monitoring [4, 5]. However, the multitude of Raman-active components coupled with the endogenous tissue fluorescence usually makes the quantitative determination of the analyte of interest difficult.

As discussed elsewhere in this thesis, multivariate calibration (MVC) provides a powerful tool for spectroscopy based chemical quantification by performing multiple measurements of the sample analytical responses. The primary role of MVC methods is to develop a calibration model connecting the measured spectral signals to specific sample properties (such as constituent concentrations) for predicting the same properties of prospective samples. Typically, MVC methods such as ordinary least squares (OLS) and partial least squares (PLS) [6, 7] employ the full spectral information to differentiate the analyte of interest from the spectral interferents. To further refine the predictive ability of the MVC techniques, appropriate selection of the spectral

data points (wavelengths) has been investigated using various optimization tools such as simulated annealing [8] and genetic algorithms [9]. In fact, it has been demonstrated that optimal predictions are obtained by selecting only the analyte-specific spectral features in order to eliminate uninformative and spurious regions from calibration [10, 11]. In this regard, Raman spectroscopy is particularly suitable for wavelength selection purposes due to its inherently narrow vibrational features.

However, the application of existing wavelength selection approaches has hitherto been largely restricted to linear calibration techniques. As shown in Chapter 5, the linear additivity assumption of the basis spectra [12] may not be valid under all circumstances necessitating the introduction of non-linear modeling. This is particularly important for biomedical applications where the sample optical properties and measurement conditions may vary substantially causing the inclusion of irrelevant sources of variation in linear calibration models. Clearly, such models – full spectrum or otherwise – are unable to perform accurate predictions in prospective samples.

This chapter investigates the applicability of wavelength selection for linear and non-linear MVC methods to further enhance the robustness and quality of the calibration models. In this work, robust implies the ability to make accurate predictions irrespective of the specific identity of the sample or subject. Using tissue phantom [13] and human subject datasets [14], we show that even with a substantial reduction in the number of wavelengths sampled, we can arrive at calibration models of equivalent prediction accuracy (i.e. less than 5% deviation from full spectrum analysis). Specifically, use of wavelength selection in conjunction with non-linear support vector regression (SVR) provides improved accuracy, in comparison to PLS full spectrum analysis. Our analysis also indicates the presence of an intrinsic subset of spectral points dictated by the vibrational responses of the specific analyte and the interferents in the

samples, and the MVC method used. This leads to a minimum allowable size of the spectral subset required to establish a clinically accurate model.

Furthermore, we demonstrate, for the first time to the best of our knowledge, the prospective transferability of *a selected spectral subset* across different human subjects while maintaining equivalent levels of prediction accuracy. This is a significant result from a calibration maintenance and transfer perspective that opens avenues for robust application of spectroscopic algorithms across a larger human population and over longer periods of time. Our results here indicate the development of a wavelength selected support vector regression provides a promising calibration platform for various spectroscopic applications including diagnosis of other diseases (e.g. cancer and atherosclerosis) as well as pharmaceutical and forensic analysis.

Another goal of this chapter is to employ the results of our wavelength selection studies in designing a miniaturized Raman system for biomedical applications. Despite the substantive progress in biomedical Raman spectroscopy, its clinical translation has been impeded by two major drawbacks, namely long acquisition times and large spatial footprint of the conventional instruments. Significant research has been carried out over the past decade to overcome these drawbacks. In particular, the recent developments of lab-on-chip surface enhanced resonance Raman (SERS) systems provide a promising solution to the aforementioned issues for analysis of biochemical mixtures [15, 16]. However, the lack of robust affordable SERS substrates and the necessity for additional bio-compatible tags presents considerable difficulties for non-invasive and quantitative applications [17]. In order to preserve the intrinsic advantages of spontaneous Raman spectroscopy, multiple research groups, including our own laboratory, have designed portable systems for clinical applications [18, 19], which are optimized for collection efficiency

and anatomic accessibility. Nevertheless, a majority of the existing systems are still too large for ready application in a clinical setting, as well as for point-of-care diagnostics and personal use.

To overcome this problem, Vo-Dinh and co-workers have proposed a novel system architecture employing acousto-optic tunable filters (AOTF) and avalanche photodiodes (APD) to develop hand-held Raman spectrometers [20, 21]. While this greatly alleviates the bottleneck associated with spatial footprint of conventional systems, the polarization sensitivity of typical AOTF components hinders the collection of (polarization scrambled) Raman photons from the sample. Here, we propose an alternate design embodiment, namely tunable excitation-single detection, which eliminates the conventional dispersive spectrograph-CCD combination by using serial acquisition of the sampled wavelengths. Such a design framework, coupled with the aforementioned wavelength selection, provides not only a reduction in spatial footprint but also of the requisite acquisition time. We anticipate that the development of this system will substantially bridge the gap between routine clinical monitoring and vibrational spectroscopy based diagnosis in the immediate future.

8.2 Theoretical formulation: Wavelength selection

Over the years, wavelength selection techniques have been demonstrated (theoretically [22, 23] and experimentally [24]) to improve calibration model accuracy and robustness. Potentially, wavelength selection can enhance the stability of the model to the collinearity in the acquired Raman spectra as well as increase the interpretability of the relationship between the model and the sample compositions by reducing the number of loading vectors to the chemical rank of the system [25]. The important consideration is how many and which of the spectral bands facilitate accurate prediction. Clearly, such selection is also dependent on the rest of the constituents in the

investigated sample(s), because of the spectral overlap which maybe observed in certain fingerprint regions.

For orientation, a brief summary of the existing wavelength selection methods is provided here. Although the wavelength selection methods are sufficiently general to work in conjunction with any calibration scheme, most of the work reported in the literature is on ordinary least squares (OLS) and partial least squares (PLS) analysis. A number of wavelength selection procedures primarily differ from one another in the objective criterion used for measuring the optimality of selected subsets or in the search algorithm for the determination of these subsets. Algorithms, such as simulated annealing (SA) and genetic algorithms (GA), have been proposed as global optimizers capable of determining the best set of parameters and selecting well-defined spectral regions instead of single data points scattered across the spectral range. These stochastic search methods accept transitory reductions of predictive quality during an optimization procedure enabling them to escape local extrema without supervision. However, this stochastic nature (involving predictable processes as well as random actions) is also a major disadvantage in establishing a universal spectral subset since it is almost impossible to recreate identical GA or SA models [26].

Alternately, one can employ spectral interval selection called moving window partial least squares regression (MWPLSR) [27-29]. MWPLSR, and its variants (changeable size moving window partial least squares and searching combination moving window partial least squares), are advantageous in searching for informative spectral regions for multi-component spectral analysis. This method applies PLS calibration models in every window which moves over the spectrum and selects informative regions on the basis of lowest sum of residual errors. Over the last decade, such a moving window approach based on minimization of the residue error has

shown promising results in improving prospective prediction of analytes in mixture solutions [30].

In the following, we use a similar scheme of selecting lower residue error wavelength regions for both linear (PLS) and non-linear (support vector regression, SVR) multivariate regression models. To implement this, we construct a spectral window of size w, which starts at i^{th} spectral channel and ends at $(i+w-1)^{th}$ spectral channel. The window is progressively moved in the full spectrum range. PLS and SVR models are built and root mean squared error of validation (RMSEV) is calculated for each spectral window position by performing prediction on the validation dataset (as detailed in section 8.3.2). Subsequently, we plot the RMSEV as a function of the position of the spectral window position. The spectral window positions with large errors imply that the responses at these spectral channels are highly contaminated by the factors that cannot be modeled accurately using the calibration samples. It should be noted that residue error plot selection method searches for spectral intervals (bands), as opposed to individual scattered points, based on the continuity of the spectral response in vibrational spectra (where the characteristic FWHM is typically at least 4-16 cm^{-1}) [27].

8.3 Materials and methods

8.3.1 Experimental

We have employed two data sets to investigate: (a) the relative prediction performance of wavelength selected linear and non-linear calibration models, and (b) the transferability of the wavelength selected subset from one sample to another. For (a), we employ a physical tissue model (tissue phantom) study, which focuses on glucose detection in a multicomponent mixture under controlled laboratory settings. To accomplish objective (b), a clinical dataset from human

subjects undergoing oral glucose tolerance tests (OGTT) is analyzed. These two data sets were originally reported in our laboratory's previous publications [13, 14].

For the tissue phantom data set, spectra were collected from 50 tissue phantoms containing randomly varying concentrations of two Raman active analytes, glucose and creatinine, between 5mM to 30mM [13]. These samples also contained randomized concentrations of India ink and intralipid to mimic the turbidity values - absorption (0.09 to 0.18 cm^{-1}) and scattering (48.4 to 95.1 cm^{-1}) - in human skin tissue. Spectroscopic measurements were performed on aliquots of these tissue phantoms in a fused silica cuvette by our bench-top Raman system (shown in Fig. 3.2(A)). For the human subject dataset, transcutaneous blood glucose measurements were performed on healthy Caucasian and Asian volunteers in our laboratory [14]. Further details of the experimental protocol can be found in Chapter 3. For the ensuing analysis, datasets from volunteers exhibiting impaired glucose tolerance profiles (due to the risk of spurious correlations with photobleached fluorescence levels) and inadequate SNR have been excluded.

8.3.2 Data Analysis

In the tissue phantom study, 20 prediction samples are randomly chosen from the entire data set and kept aside for prospective application. The creation of an independent prediction set is a conventional approach used to mitigate and/or test for the presence of spurious correlations. Subsequently, the remaining 30 tissue phantoms are randomly split into twenty calibration samples and ten validation samples. The moving window approach is used to generate a regression vector from the calibration sample spectra corresponding to the specific window position. This regression vector is then used on the validation samples to obtain a RMSEV value. The window is subsequently moved, as discussed in the Theoretical Formulation section, over the full spectrum to construct the residue error plot as a function of the moving window position.

Figure 8.1 shows a residue error plot calculated from a representative partition of calibration and validation sets. From the residue plot, spectral points with lowest computed RMSEV are selected for developing the final regression vector. For our analysis, we selected 300 to 900 spectral points (in increments of 100), where one spectral point roughly corresponds to 1.45 cm^{-1}. The final regression vector is generated from the thirty tissue phantoms constituted by the calibration and validation data sets for each subset of spectral points. This regression vector is used prospectively on the corresponding prediction set and RMSEP is calculated for the specific subset of spectral points. It is worth mentioning that the prediction set also comprises the spectral information *only* at the points selected to create the regression vector. To ensure the reproducibility of our prediction results, 100 iterations are performed to obtain an average RMSEP.

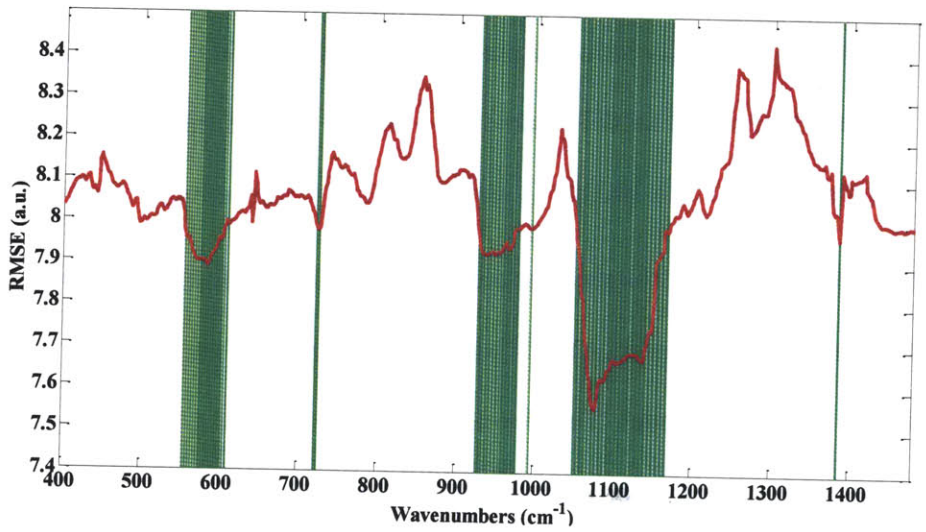


Fig. 8.1 Residue error plot (solid red curve) calculated for SVR calibration model in the tissue phantom dataset using a 20 spectral point window. The set of 300 spectral points which exhibit the minimum residue error are highlighted in green.

For our investigation, different window sizes from 10 to 20 spectral points are selected. This size range corresponds to the full width at half maximum (FWHM) of the prominent bands in the glucose Raman spectrum. Assigning window sizes beyond this range exhibited adverse effects on the resultant residue lines, as also noted by Jiang *et al.* [27]. Wavelength selection using the above protocol is performed for both PLS and SVR calibration. PLS models are created based on the number of loading vectors that provide the least error in cross-validation for full spectrum analysis. A standard recommendation for PLS calibration is to incorporate at least three times the number of samples as the rank of the calibration model. To satisfy this criterion, twenty samples are chosen for our calibration studies (as six loading vectors provided the least error in cross validation). The SVR calculations are carried out using a SVM MATLAB toolbox [31]. Prior to wavelength selection for SVR processing, the Raman spectra are linearly scaled by dividing each spectrum by the maximum intensity value to prevent the skewed effect arising from the larger pixel intensity values. The optimal model parameters C and σ^2 are obtained by employing a grid search algorithm in the range of 1 to 10000 (C) and 0.01-10 (σ^2), respectively.

Similar analysis steps are also followed for the human subject dataset. From a representative human volunteer data set (volunteer A) all but five data points are used to develop the wavelength-selected regression model. The developed regression model is then applied on the remaining five data points of volunteer A to evaluate the RMSEV as a function of the moving window position and subsequently for the construction of the residue error plots. To obtain enhanced robustness in the wavelength subset selection, we perform 100 iterations by re-partitioning the calibration and validation data sets of volunteer A. The residue error plots from these iterations are added to form a cumulative error plot as a function of wavelength. The spectral points with the least cumulative error are selected for prospective application in the other

human volunteer datasets. For clarity, we henceforth represent the remaining volunteers as volunteer B_i, where i is the index of the volunteer. For volunteer B_i, the data points are split into calibration and prediction. The regression models are constructed on the calibration data points of volunteer B_i using only the spectral points selected from volunteer A. These models are subsequently used to estimate the glucose concentrations of the prediction data points in volunteer B_i. The resultant prediction errors provide a true measure of the prospective applicability of the selected spectral subset, i.e. the transferability of the selected points across human subjects. The loading vectors for PLS analysis are optimized for each volunteer based on leave-one-out cross-validation on the calibration data points. Similar SVR optimization is also carried out for C and σ^2 in the range of 1 to 10000 and 0.01 to 10, respectively.

It is worth mentioning that for both the human volunteer as well as the tissue phantom study, the spectra were used directly for development of the MVC calibration models without any data manipulation such as removal of fluorescence.

8.4 Results and discussion

8.4.1 Comparison of wavelength selection in PLS and SVR in tissue phantom studies

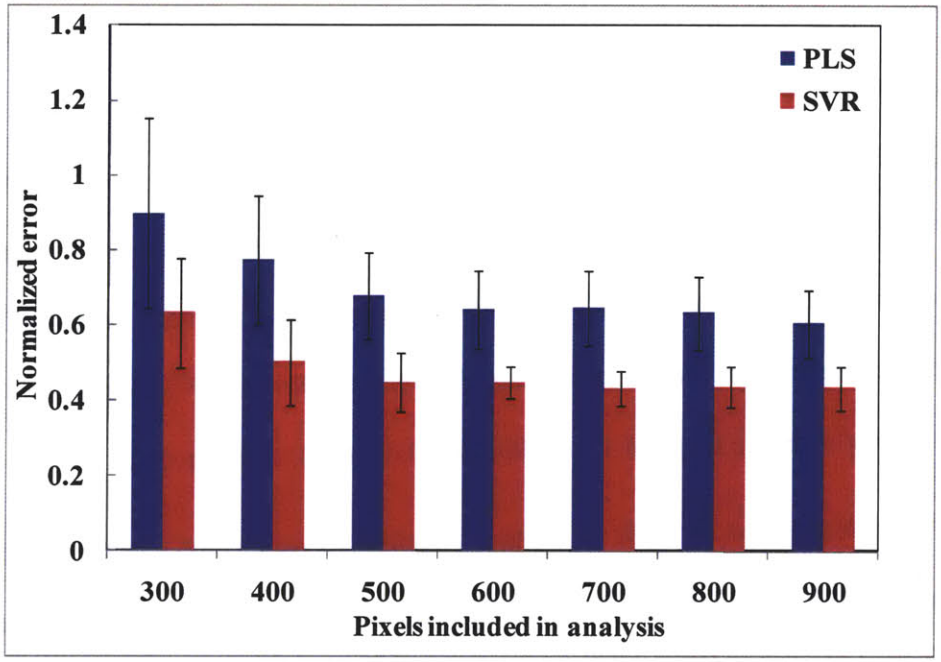


Fig. 8.2 Bar plot showing comparative performance of wavelength-selected PLS (blue) and SVR (red) calibration models in tissue phantoms for glucose prediction. The lengths of the bars are proportional to the average RMSEP and the associated error bars represent the standard deviation of the RMSEP over 100 iterations.

Wavelength selection was performed on the tissue phantom dataset using both PLS and SVR models. Figure 8.2 shows the results of prospective prediction for glucose obtained with PLS (blue) and SVR (red) models, where the lengths of the bars are proportional to the average RMSEP and the associated error bars represent the standard deviation of the RMSEP over 100 iterations. The figure provides a comparative estimate of predictive performance of calibration models corresponding to the selection of minimum residue wavelength subsets of different sizes, ranging from 300 to 900 spectral points. It is evident that the SVR calibration models always

outperform the PLS calibration models with identical size of wavelength subset in regard to prospective prediction. For example, given a spectral subset size of 300 points, the mean prediction errors for glucose are observed to be 0.89 mM and 0.63 mM for the PLS and SVR models, respectively. In fact, the SVR model with 300 spectral points provides equivalent levels of prediction accuracy as the full spectrum PLS model (mean RMSEP of 0.6 mM).

In conjunction with RMSEP determination, we have also evaluated the relative predictive determinant (RPD) metric to classify the overall prediction quality of the individual calibration models for the two Raman-active analytes (glucose and creatinine). Briefly, RPD is defined as the ratio of the standard deviation of the reference concentration in the sample population (σ_R) to the standard error of prediction (the standard deviation of the differences between predicted and reference values) (σ_{R-P}):

$$RPD = \frac{\sigma_R}{\sigma_{R-P}} \tag{8.1}$$

Typically, a RPD value of 5 is considered to be good for quality control while a value larger than 6.5 is acceptable for process monitoring. A calibration model, with a RPD value higher than 8, may be used for any application. From Table 8.1, we observe that both PLS and SVR models including at least 300 spectral points show excellent prediction quality. To this end, RPD provides a useful tool in evaluating the allowable minimum size of the spectral subset for different applications and analytes of interest [32].

Table 8.1 Summary of statistics obtained from relative predictive determinant analysis for glucose and creatinine using PLS and SVR calibration

	PLS			SVR		
	100 pixels	300 pixels	Full spectrum	100 pixels	300 pixels	Full spectrum
RPD (glucose)	3.9	8.45	12.4	1.6	12.2	16.9
RPD (creatinine)	5.9	11.4	17.1	2.4	15.2	24.2

From Fig. 8.2, we observe that, as the spectral subset size is decreased initially from 900 to 500 spectral points, the prediction errors are not substantially altered. However, with further decrease of the size of the spectral subset, the prediction errors rise considerably. This threshold measure of the spectral subset size provides an idea of relevant information specific to the analyte of interest present in the full spectrum. Particularly for our tissue phantom study, it would appear that the glucose-specific informative regions constitute less than half of the full spectrum. Presumably, wavelength selection also enhances the stability of the model to the collinearity in the acquired Raman spectra and increases the interpretability of the relationship between the model and the sample compositions by reducing the number of loading vectors to the chemical rank of the system.

8.4.2 Comparison of wavelength selection in PLS and SVR in human subject studies

From our tissue phantom studies, we observe that reduction of spectral points (even by a factor of 3) does not substantially deteriorate calibration model performance, as long as the appropriate wavelengths depending on the analyte of interest and spectral interferents are analyzed.

Nevertheless, the prospective transfer of the selected wavelength subset from one human to another is substantially more difficult because of the complexity arising from substantive variations in tissue optical (e.g. turbidity, autofluorescence and skin heterogeneity) and physiological properties (dynamics of the analyte of interest) and changes in experimental conditions. Indeed, while several investigators have successfully applied wavelength selection in powder and mixture samples, to the best of our knowledge its transferability in complex biological specimens has not been previously demonstrated.

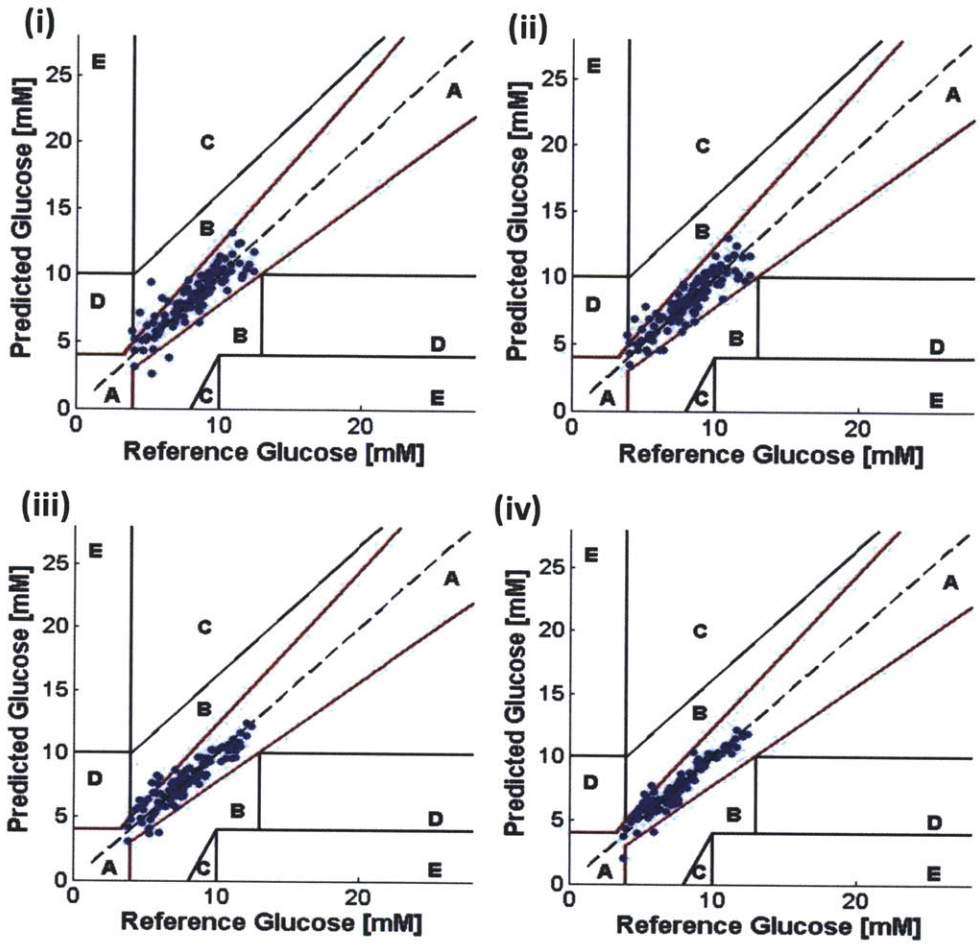


Fig. 8.3 Glucose prediction performance of wavelength-selected PLS and SVR calibration models in human subjects shown on the Clarke error Grid. PLS calibration results are shown in (i) and (ii) for 300 and 900 spectral points respectively. SVR calibration results are shown in (iii) and (iv) for 300 and 900 spectral points respectively.

Figure 8.3 shows the prediction results of PLS and SVR calibration models using wavelength selection on the human subject dataset. It is worth emphasizing that the wavelength subsets were obtained using the residue error plot method in one individual and the same wavelength subset was prospectively applied on all remaining subjects with no further analysis or alteration. The prediction results are shown plotted on Clarke error grids, which are widely used for quantifying the clinical accuracy of blood glucose monitors. Predictions in zones A and B are regarded as

acceptable while those in zones C, D and E are potentially dangerous if used for determining treatment options. The RMSEP and the R^2 correlation-coefficient values in each of the four cases (i.e. for PLS and SVR calibration with 300 and 900 spectral points, respectively) is provided in Table 8.2. We observe that the average RMSEP for SVR in each case is much lower than the corresponding value for PLS, such that the SVR calibration model with 300 spectral points provides equivalent (or better) prediction accuracy as compared to the PLS 900 spectral point model.

Table 8.2 Comparison of wavelength selection for PLS and SVR calibration models in human subjects

	PLS		SVR	
	300 spectral points	900 spectral points	300 spectral points	900 spectral points
RMSEP	18.6	16.9	15.1	11.3
R^2	0.87	0.89	0.92	0.95
Percentage of points satisfying ISO criteria	86.66	89.17	88.33	94.17

Current FDA recommendations (ISO 15197 guideline) stipulates that for clinical usage 95% of the sensor predictions should be within 15 mg/dL (0.83mM) of reference for glucose <75 mg/dL (4.2 mM) and within 20% for glucose ≥75 mg/dL (4.2 mM) (as discussed in Chapter 2). The percentage of data points satisfying the FDA criteria for these four calibration models are tabulated in Table 8.2. While none of the models completely satisfy the 95% criterion (even

though the SVR model for 900 spectral points is very close), one can reasonably expect that the deviations from the aforementioned criterion can be overcome by correcting for variations in tissue turbidity and autofluorescence as well as by addressing the physiological lag between blood and interstitial fluid glucose. Furthermore, given the increase in spectral SNR that can be obtained by additional refinement of the instrumentation including incorporation of non-imaging optics [33] it may be possible to develop clinically accurate SVR models employing less than 300 spectral points.

Importantly, our results reflect that the selected wavelength subsets are transferable from one human subject to another. To quantify this notion of transferability, we compare the degree of overlap between the wavelength subsets giving the minimum residue error for each of the human subjects. Evidently, if the degree of overlap between wavelength subsets from two human subjects is high, prospective application of the selected wavelength subset from one subject to the other will also provide accurate results. Here, we define robustness metric as the ratio of overlapping spectral points (between selected wavelength subsets of any two volunteers) to the total number of selected spectral points (300). Figure 8.4 provides box-plots of the robustness metrics for PLS and SVR calibration. The box-plots are generated from the robustness metrics evaluated for all possible combinations in the human subject dataset. Clearly, SVR provides a marked improvement with respect to transferability of the selected wavelength subsets in relation to PLS. This enhancement opens new avenues toward construction of universal calibration models based on a small set of features, which also enables the development of Raman instruments employing a smaller set of wavelength sampling channels.

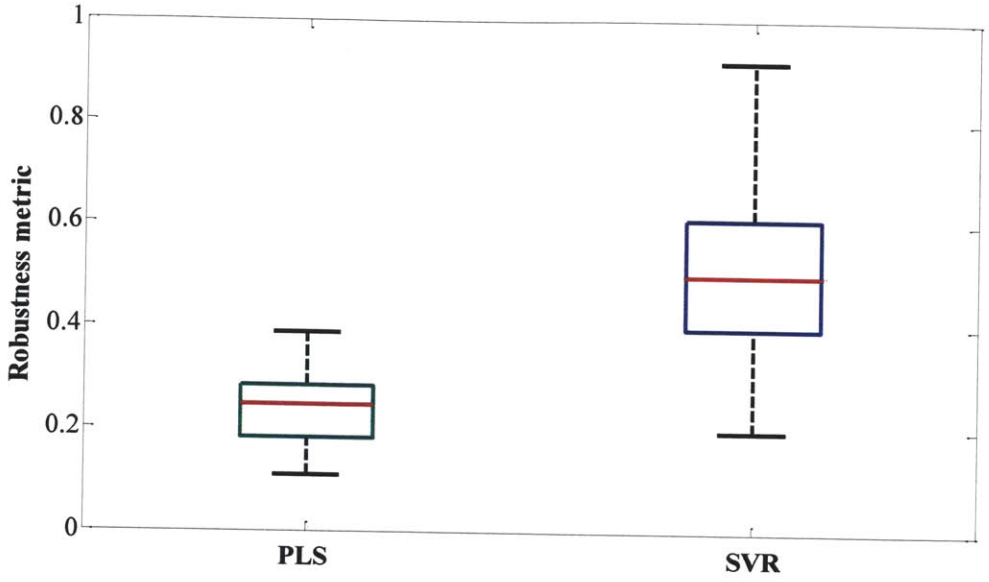


Fig. 8.4 Boxplot showing robustness metric is for PLS and SVR calibration models in prospective prediction of human subject data.

8.4.3 Design of miniaturized Raman instrument

As mentioned in the Introduction, a major drawback in widespread translation of Raman spectroscopy is the large spatial footprint of the conventional setups, which is ill-suited to meet the needs of a clinical setting. Specifically, continuous glucose monitoring necessitates the development of a hand-held or a wearable device due to the frequency of measurements required for diabetic patients. In this regard, the use of fiber probe-based instruments has considerably eased the problem because of the additional flexibility and anatomic accessibility provided by such systems in comparison to conventional free-space optical bench setups [18]. Nevertheless, the bottleneck associated with construction of a truly handheld device emanates from the large spatial footprint of the conventional dispersive spectrograph – CCD combination. Here, we investigate the feasibility of introducing miniaturized Raman instruments that can provide

adequate detection sensitivity and the effect of performing wavelength selection on the implementation of such designs.

The simplest strategy for substituting the dispersive spectrograph-CCD combination is by employing a set of optical bandpass filters with corresponding photo-detectors. However, this approach is feasible only when a limited number of bands (spectral points) are sufficient to provide clinical accuracy in spectroscopic predictions. In multi-component mixtures, such as in biological samples, this is rarely possible, e.g. reduction of sampled spectral points below 300, results in a rise in glucose concentration prediction error in the human subject dataset. On the other hand, the application of tunable filters, such as liquid crystal tunable filter (LCTF) and AOTF [34], circumvents the need for individual bandpass filters by electronically modulating the transmission band of the filters over a broad spectral range. Furthermore, the use of tunable filters facilitates the application of the same instrument for multiple diagnostic applications since the selection of wavelengths does not necessitate a change in the system hardware.

Due to the intrinsic advantages of rapid tunability and precise control, the aforementioned filters can be advantageously employed in two analogous designs for a miniaturized Raman clinical device: (A) fixed laser line excitation with tunable detection (as suggested by Vo-Dinh and co-workers) [20, 21] and (B) tunable laser excitation ('sweeping source') with fixed wavelength detection, which remains largely untested for spontaneous Raman spectroscopy applications. These designs are analogous in the sense that changing either the excitation or detection wavelength while keeping the other constant results in acquisition of identical Raman bands (as Raman shift is a measure of the change in photon emission energy from that of the excitation source). The major difference between these designs and the conventional spectrometer-CCD systems is that while the latter can be used to acquire spectral information at

all wavelengths simultaneously, these designs only allow serial sampling of the wavelengths. However, when the filters are used in conjunction with an avalanche photo-diode (APD) or a photomultiplier tube (PMT), which provide significant signal amplification, the serial acquisition process can be implemented in as short a time frame as is employed for conventional spectrometers.

In principle, the combination of AOTF and APD (design (A)), is a strategy that can be implemented to overcome the footprint drawbacks. While this design has provided promising results in chemical analysis, the polarization sensitivity of commercially available AOTF systems limit the collection efficiency of the Raman photons from the sample, due to the scrambled nature of the back-scattered light from the sample. Although it is possible to employ a polarization rectification unit consisting of a beam displacement prism and a half-wave plate, this new unit contributes an additional layer of complexity and size to the overall system. To overcome this problem, we propose the sweeping source design, which does not require a tunable filter on the detection side.

For the sweeping source design, one can employ a variety of wavelength tunable sources that have been proposed in the literature [34, 35]. Specifically, wavelength sweeping can be performed by inserting a combination of dispersive optical elements (such as gratings) and mechanical components (e.g. rotating polygonal mirror) inside of the laser cavity. In addition, it can be implemented based on tunable filters, such as AOTF and Fabry-Perot filters. For our system, we propose to employ a semiconductor optical amplifier in conjunction with AOTF to attain the appropriate tunability, while retaining a small footprint. As shown in Fig. 8.5, the excitation light from the sweeping source is coupled through a focusing lens to the excitation fiber. The fiber probe is constructed by positioning multiple collection fibers around the central

excitation fiber. The back-scattered light is collected via the collection fibers and transmitted through a collimating lens to a bandpass filter that transmits only the requisite spectral bands. The transmitted photons are detected by the APD. A photo-multiplier tube (PMT) can also be used in place of the APD, especially because of the significantly higher internal gain, but its efficiency in the NIR region impedes its application. The primary advantage of this design over the conventional designs stems from the significantly smaller spatial footprint requirements. The proposed design can be readily implemented in a hand-held box and sets the stage for even smaller systems given the current developments in the area of MEMS-based tunable sources [36].

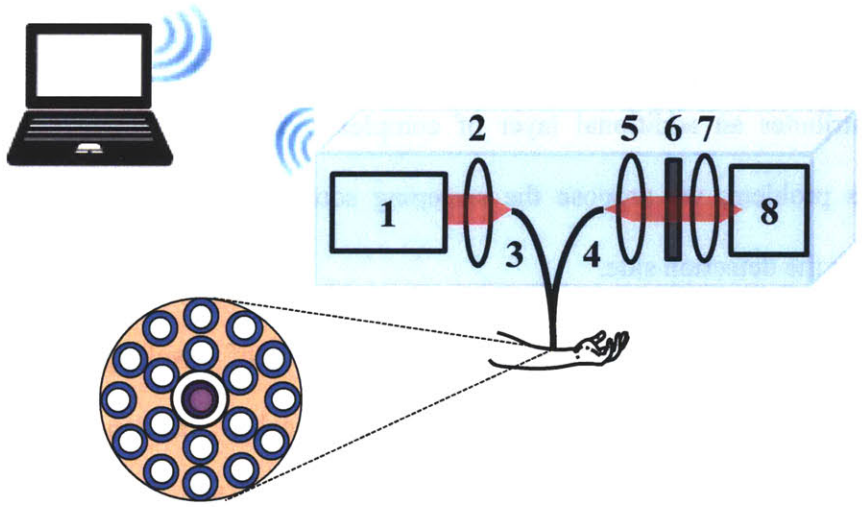


Fig. 8.5 Schematic of the proposed handheld device. Here, component 1 is the tunable laser source; 2, 5 and 7 represent focusing and collimating lenses; 3 represents the excitation fiber; 4 represents the collection fiber; 6 indicates the band pass filter; 8 is the photo-detector.

It is important to explore the effect of our wavelength selection results on the operating parameters of the proposed Raman systems (design (A) and (B)). Since these designs for the

miniaturized systems employ serial acquisition of the photons at different wavelengths, a significant number of the Raman photons are not utilized in constructing the final spectrum. In addition, the photo-detector amplification does not improve the analyte signal-to-noise ratio, which can only be improved by increasing acquisition time. While this problem cannot be completely eliminated, appropriate application of wavelength selection can greatly alleviate the problems associated with serial acquisition.

For the specific problem of transcutaneous glucose detection, we have demonstrated that a set of 300 spectral points can provide equivalent levels of accuracy as the full spectrum (900 spectral points). This implies that instead of having to perform serial acquisition over 900 spectral points, one can acquire the Raman photons at one-third of the spectrum, which in turn means the total acquisition time can be reduced by a factor of three. This result is of great significance especially in biomedical applications where acquisition time is limited by clinical constraints (e.g. spectral diagnosis in an operating room). Conversely (and more importantly for glucose detection), one can acquire for longer periods of time (three-fold) at the appropriate wavelengths that leads to more efficient utilization of the important Raman photons. Considering that our detection is limited by shot-noise, a three-fold increase in acquisition time translates to a 1.73 times increase in analyte SNR. Using the minimum detectible concentration formulation [37], the increase in SNR results in a corresponding reduction in the prediction uncertainty (or rise in precision) by nearly 43%.

8.5 Summary

In this chapter, we have employed wavelength selection using linear (PLS) and non-linear calibration (SVR) models. We have demonstrated that the prediction accuracy is substantially improved by using SVR. In fact, our studies indicate that SVR models can provide the same

prediction accuracy with a small fraction of the spectral information as used in PLS full spectrum analysis. Projecting our results to biological applications may result in cheaper smaller sized systems that employ simpler optics for imaging of a select set of bands. Additionally, we have proposed a novel sweeping source design for developing a miniaturized Raman instrument that greatly reduces the spatial footprint constraints imposed by the conventional spectrograph-CCD combination. Furthermore, it is shown that the application of wavelength selection to this sweeping source design (or to an analogous tunable detector scheme) considerably reduces the acquisition time. This can conversely be employed to increase analyte SNR over the same period of acquisition resulting in better predictive ability.

We anticipate that such a transformation may help bridge the gap between routine glucose monitoring and molecular spectroscopy in the immediate future and in a manner that is useful to clinical laboratories as well as individual patients. Additionally, the approach proposed here is sufficiently broad and general for the development of handheld diagnostic devices that address other major disease classes including cancer diagnosis in breast and mucosal tissue.

8.6 References

(1) Frank, C. J.; Redd, D. C. B.; Gansler, T. S.; McCreery, R. L. *Analytical Chemistry* **1994**, 66, 319-326.

(2) Haka, A. S.; Shafer-Peltier, K. E.; Fitzmaurice, M.; Crowe, J.; Dasari R. R.; Feld, M. S. *Proc. Natl. Acad. Sci.* **2005**, 102, 12371-12376.

(3) Motz, J. T.; Fitzmaurice, M.; Miller, A.; Gandhi, S. J.; Haka, A. S.; Galindo, L. H.; Dasari, R. R.; Kramer J. R.; Feld, M. S. *Journal of Biomedical Optics* **2006**, 11, 021003-9.

(4) Enejder, A. M. K.; Koo, T. W.; Oh, J.; Hunter, M.; Sasic, S.; Feld M. S.; Horowitz, G. L.; *Optics Letters* **2002**, 27, 2004-2006.

(5) Chaiken, J.; Finney, W.; Knudson, P. E.; Weinstock, R. S.; Khan, M.; Bussjager, R. J.; Hagrman, D.; Hagrman, P.; Zhao, Y. W.; Peterson, C. M.; Peterson, K. *Journal of Biomedical Optics* **2005**, 10, 031111.

(6) Martens, H.; Naes, T. *Multivariate Calibration*, Wiley, New York, **1989**.

(7) Haaland, D. M.; Thomas, E. V. *Analytical Chemistry* **1988**, 60, 1193-1202.

(8) Kalivas, J. H.; Roberts, N.; Sutter, J. M. *Analytical Chemistry* **1989**, 61, 2024-2030.

(9) Leardi, R. *Journal of Chemometrics* **2000**, 14, 643-655.

(10) Liang, Y.; Xie, Y.; Yu, R. *Analytica Chimica Acta* **1989**, 222, 347-357.

(11) Ding, Q.; Small, G. W.; Arnold, M. A. *Analytical Chemistry* **1998**, 70, 4472-4479.

(12) Hanlon, E. B.; Manoharan, R.; Koo, T. W.; Shafer, K. E.; Motz, J. T.; Fitzmaurice, M.; Kramer, J. R.; Itzkan, I.; Dasari, R. R.; Feld, M. S. *Physics in Medicine and Biology* **2000**, 45, R1-R59.

(13) Bechtel, K. L.; Shih, W. C.; Feld, M. S. *Optics Express* **2008**, 16, 12737-12745.

(14) Enejder, A. M. K.; Scecina, T. G.; Oh, J.; Hunter, M.; Shih, W.; Sasic, S.; Horowitz, G. L.; Feld, M. S. *Journal of Biomedical Optics* **2005**, 10, 031114-9.

(15) Quang, L. X.; Lim, C.; Seong, G. H.; Choo, J.; Do, K. J.; Yoo, S. *Lab on a Chip*, **2008**, 8, 2214-2219.

(16) Lim, C.; Hong, J.; Chung, B. G.; deMello, A. J.; Choo, J. *The Analyst* **2010**, 135, 837-844.

(17) Stoddart P.; White, D. *Analytical and Bioanalytical Chemistry*, **2009**, 394, 1761-1774.

(18) Motz, J. T.; Gandhi, S. J.; Scepanovic, O. R.; Haka, A. S.; Kramer, J. R.; Dasari, R. R.; Feld, M. S. *Journal of Biomedical Optics* **2005**, 10, 031113-7.

(19) Lieber C.; Mahadevan-Jansen, A. *Optics Express* **2007**, 15, 11874-11882.

(20) Cullum, B. M.; Mobley, J.; Chi, Z.; Stokes, D. L.; Miller, G. H.; Vo-Dinh, T. *Review of Scientific Instruments* **2000**, 71, 1602-1607.

(21) Mobley, J.; Cullum, B. M.; Wintenberg, A. L.; Frank, S. S.; Maples, R. A.; Stokes, D. L.; Vo-Dinh, T. *Review of Scientific Instruments* **2004**, 75, 2016-2023.

(22) Xu, L.; Schechter, I. *Analytical Chemistry* **1996**, 68, 2392-2400.

(23) Spiegelman, C. H.; McShane, M. J.; Goetz, M. J.; Motamedi, M.; Yue, Q. L.; Cote, G. L. *Analytical Chemistry* **1998**, 70, 35-44.

(24) Jouan-Rimbaud, D.; Walczak, B.; Massart, D. L.; Last I. R.; Prebble, K. A. *Analytica Chimica Acta* **1995**, 304, 285-295.

(25) Soares, I. P.; Rezende, T. F.; Silva, R. C.; Castro E. V. R.; Fortes, I. C. P. *Energy & Fuels* **2008**, 22, 2079-2083.

(26) Igne B.; Hurburgh, J. C. R. *Journal of Chemometrics* **2010**, 24, 75-86.

(27) Jiang, J.; Berry, R. J.; Siesler, H. W.; Ozaki, Y. *Analytical Chemistry* **2002**, 74, 3555-3565.

(28) Kasemsumran, S.; Du, Y. P.; Murayama, K.; Huehne, M.; Ozaki, Y. *Analytica Chimica Acta* **2004**, 512, 223-230.

(29) Du, Y. P.; Liang, Y. Z.; Jiang, J. H.; Berry, R. J.; Ozaki, Y. *Analytica Chimica Acta* **2004**, 501, 183-191.

(30) Sahin, S.; Sarburun, E.; Demir, C. *Analytical Methods* **2009**, 1, 208-214.

(31) Chang, C. C.; Lin, C.-J. LIBSVM: a library for support vector machines, **2001**.

(32) Williams, P. *Near-infrared technology in the agricultural and food industries*, 2nd edition, P. Williams, K. Norris (eds.), AACC Inc., St. Paul, MN, USA, **2001**.

(33) Tanaka, K.; Pacheco, M. T. T.; Brennan III, J. F.; Itzkan, I.; Berger, A. J.; Dasari, R. R.; Feld, M. S. *Applied Optics* **1996**, 35, 758-763.

(34) Harris, S. E.; Wallace, R. W. *Journal of the Optical Society of America* **1969**, 59, 744-747.

(35) Masterson, H. J.; Sharp, G. D.; Johnson, K. M. *Optics Letters* **1989**, 14, 1249-1251.

(36) Datta, M.; Pruessner, M. W.; Kelly, D. P.; Ghodssi, R. *Solid-State Electronics* **2004**, 48, 1959-1963.

(37) Scepanovic, O. R.; Bechtel, K. L.; Haka, A. S.; Shih, W.; Koo, T.; Berger, A. J.; Feld, M. S. *Journal of Biomedical Optics* **2007**, 12, 064012-10.

CHAPTER 9

CONCLUSIONS AND FUTURE DIRECTIONS

Blood analytes provide valuable information for the diagnosis of many diseases and related health conditions. Development of painless, non-invasive methods for measuring such analytes has been extensively investigated. In particular, non-invasive blood glucose monitoring has received considerable attention due to its important implications for diabetes management and therapeutics. Various techniques ranging from electrochemical assays to optical methods have been proposed to meet the goals of painless and accurate blood glucose measurements.

In this context, NIR Raman spectroscopy has shown substantial promise and provided successful predictions of glucose at physiologically relevant concentrations in serum, whole blood, and even in individual human subjects. However, successful clinical implementation over longer periods of time or across a wider cross-section of population has not been demonstrated. Consequently, researchers have tried to identify the specific factors (root-causes) that degrade the glucose measurement accuracy of Raman spectroscopy by introducing non-analyte specific variance into the calibration model. The predominant factors include sample-to-sample variability in absorption and scattering properties (turbidity), tissue autofluorescence and associated quenching, and physiological lag between blood and interstitial fluid (ISF) glucose. Clearly, a methodological implementation and validation from bench to bedside is required to make non-invasive glucose measurement a viable clinical technology.

9.1 Overview of thesis accomplishments

The goal of this thesis was to investigate the aforementioned scientific questions and phenomena that introduce (non-analyte specific) confounding factors into this important diagnostic problem. Additionally, as discussed in Chapter 1, this thesis seeks to develop methodologies for solving the principal issues that prevent the reduction of this technology to a viable one for routine clinical diagnosis.

Specifically, Chapter 4 focused on overcoming the effects of sample-to-sample turbidity induced variations in sampling volume as well as in the acquired Raman spectra. To account for these variations, a novel method, called turbidity corrected Raman spectroscopy (TCRS), has been developed. This method is based on the photon migration approach and employs alternate acquisition of Raman and diffuse reflectance spectra. It has been demonstrated that the proposed method can extract intrinsic line shapes and intensity information from Raman spectra acquired in a turbid medium. Importantly, it has been shown to provide significant improvement in prediction accuracy over uncorrected spectral analysis, when used in conjunction with standard multivariate calibration schemes.

Chapter 5 explored the validity of the assumption that there exists a linear relationship between the acquired spontaneous Raman spectra and the reference concentrations of the analyte of interest. While this assumption is valid under most circumstances (stemming from the incoherent signal addition from the constituent molecules), it is shown that fluctuations in system and sample variables may play a major role in causing non-linearity in the regression process. In other words, if a linear regression model is constructed based on the spectra-concentration dataset, it risks the inclusion of unnecessary and spurious sources of variance and noise to explain the curved effects. We have demonstrated that tissue turbidity variation alone, especially

at high levels of absorption, can cause the deviation of the predicted to reference concentration from a linear function. To account for such curved effects, a relatively new method called support vector regression is employed that has the additional advantage of producing a reproducible global calibration model (in contrast to neural networks, for example).

In Chapter 6, the ubiquitous problem in Raman spectroscopy, namely that of fluorescence, is discussed. While the ability to re-construct intrinsic Raman signals from fluorescing mixtures is important to several applications, the two biggest hurdles for Raman-based analyte monitoring is the associated noise (both shot noise and detector noise) and the intensity variations over time, especially as a result of quenching from sustained laser exposure. In this chapter, it is established that the prediction accuracy of a calibration model can be severely compromised by the photobleaching behavior due to the introduction of spurious correlations between analyte concentrations and fluorescence levels. This is a significant result for applications where any the concentration dataset may have some form of inbuilt temporal correlation, e.g. tolerance test-based studies on diabetic patients. However, such a situation can be significantly reduced by the application of a suitable fluorescence removal strategy, which does not compromise the integrity of the Raman spectral features. In our studies, we have determined that shifted subtracted Raman spectroscopy (SSRS) presents such a tool and, additionally, is capable of removing the detector noise that represents the main source of noise at high signal levels.

Chapter 7 addressed the question of accurate reference concentrations for spectroscopy-specific measurements, given the limited penetration depth of light in skin tissue and the small volume fraction of blood vessels in the dermis. The slow diffusion kinetics for glucose between the blood and interstitial fluid compartments means that at times of relatively rapid changes of glucose levels there exists a finite differential between the two values. This creates a problem for

nearly all existing non-invasive and minimally invasive techniques, which predominantly probe interstitial fluid glucose, thereby creating inconsistencies in calibration, where blood glucose measurements are used as reference. To overcome this problem, we have introduced a dynamic concentration correction scheme, which when used in conjunction with standard spectroscopic measurements, provides substantial enhancement in prediction accuracy. In particular, the resulting improvement in blood glucose estimates should considerably enhance the spectroscopic ability to correctly determine hypoglycemia and even predict impending hypoglycemia based on the rate of change in glucose concentrations.

Finally, Chapter 8 presented a new direction in our research program, where the design and potential development of miniaturized Raman systems specialized for blood glucose monitoring is investigated. The key concept to the successful implementation of such a design lies in the reliable determination of a limited set of specific wavelengths that can provide similar glucose information content as the whole fingerprint region. Using a previously developed moving window approach, it is demonstrated that one can obtain a set of wavelengths (which is based on the analyte of interest as well as the spectral interferents present in the sample) and transfer it with adequate robustness to other similar samples. This opens up avenues for newer design constructs that can replace the conventional laser-spectrograph-CCD system while using only a fraction of the spatial footprint of the latter. Based on our current wavelength selection results, we have designed a tunable excitation based band pass detection system employing fiber probes that can be assembled in a laptop-size module. Ultimately, the goal is to build an even simpler (and cheaper) system potentially employing LED and photodetectors, which would then obviate the need for having a "universal" calibration algorithm as such devices could use person-specific calibration protocols.

9.2 Future directions

Based on the improved methodologies presented in this thesis, our laboratory has already started the first phase of clinical feasibility studies. A portable clinical instrument, based on similar clinical instruments developed for breast cancer [1, 2] and atherosclerosis diagnosis [3], has been designed and fabricated for these feasibility studies on healthy human volunteers. However, there are a few notable changes in this version of the system to maximize the glucose SNR.

Traditionally, the light is delivered to and collected from the tissue surface by a specially designed Raman fiber probe, with a central excitation fiber surrounded by annular ring(s) of collection fibers. A special shell-and-tube miniaturized optical filter assembly, consisting of a central short pass filter surrounded by a notch filter, is installed at the tip of the fiber probe to suitably reject the additional signal generated by the probe itself. The number and diameter of fibers is determined based on the conservation of étendue and the derived angular and radial collection efficiencies in the turbid media. For further details, the reader is referred to a couple of excellent reviews on the subject of probe development in biomedical optics [4, 5].

Nevertheless, only a small fraction (~10%) of the light reflected from the surface is typically collected because of the limited collection angle of the collection fibers (as determined by the numerical aperture (NA)). To overcome this persistent problem, non-imaging optical elements, such as the compound parabolic concentrators [6, 7], are incorporated into the system. These concentrators are used in reverse as angle transformers that are capable of collecting the whole hemispherical light distribution and converting them to a beam of smaller NA for efficient fiber coupling. Variations of the design, e.g. coupling of the concentrator with a lens, further serve to enhance the intrinsic ability to collect a wide range of angles, even for skew rays.

This concentrator-based instrument is currently being employed for glucose tolerance test based clinical feasibility studies on 50 healthy volunteers at the MIT Clinical Research Center. We expect that these studies will provide a definitive understanding of the improvements produced by the methodologies described in this thesis as well as the aforementioned changes in optical system design. Evidently, a significant portion of our future directions - including clinical validation studies - will be based on the results of the current feasibility investigations. However, in addition to any such new avenues, our research program in the future can be classified into two main classes, namely scientific directions and clinical directions. In the following, we outline the specific goals of both the scientific and clinical research blocks.

Scientific directions

1. Determination of optimal measurement site

A spectroscopic study should be conducted to select the optimal anatomic measurement site from which to collect Raman spectra for non-invasive blood analyte studies. As a reference point, three potential measurement sites, namely the forearm, the thenar eminence of the palm and the earlobe, may be tested. These sites, based on physiological and patient compliance considerations, have been previously selected by other researchers based on their optical properties, in order to maximize the signal-to-noise ratio of the spectra obtained from these sites [8]. Ideally, the optimal site should have a thin layer of epidermis, low pigmentation and a substantially high distribution of blood capillaries to minimize the problem of physiological glucose lag.

The chosen measurement site should be minimized for spectral variations because the identical measurement spot cannot be practically reproduced even in single-subject studies.

Spectral variations, even within each site, can arise from turbidity variations and changes in the fluorescence background. TCRS (Chapter-4) has been developed to correct for turbidity variations, which alter sampling depth, and SSRS (Chapter-6) has been introduced to alleviate the tissue autofluorescence and photobleaching issues. Nevertheless, the smaller the extent of these spectral variations, the more accurate the resulting concentration measurements are likely to be. In general, we expect that person-to-person spectral variation for a given site is larger than the variation associated with multiple measurements at the same site in a given person. As a consequence, we expect that a multi-person spectroscopic study (similar to our current clinical feasibility studies but incorporating spectral measurements from different sites) will be sufficient to determine the site showing minimum spectral variability.

2. Optimization of excitation-collection geometry for specific measurement site

Most of the optics and spectroscopy experiments for transcutaneous glucose detection till date have employed the back-scattering geometry. This can be largely attributed to the fact that given the turbid nature of biological tissue very few photons, and naturally even fewer Raman photons, can be acquired in the transmission geometry. This is particularly true if such measurements are made on relatively thicker anatomical sites such as the forearm, palm or fingertip. Nevertheless, transmission Raman experiments could be performed on a few sites such as the thenar eminence, earlobe and possibly the inner lip. Clearly, the optimization of excitation-collection geometry is intrinsically coupled to the previous scientific direction, i.e. determination of optimal measurement site, and this aspect is further emphasized in the following paragraphs.

The primary advantage to performing transmission Raman experiments is that the distribution of Raman photons collected has an approximately Gaussian profile with a peak at the center of the tissue sample. This is in contrast to the Raman experiments performed in back-

scattering geometry, where the acquired Raman photons exhibit an approximately exponential decrease with depth from the reflection surface. This is significant for glucose measurements because the epidermis (which may extend up to 100-200 micron below the surface) is completely avascular and provides no glucose information, thereby leading to a large non-analyte specific Raman signal in the reflection geometry. On the other hand, in the transmission geometry, most of the signals emanate from the region of the dermis which has a 45% volume fraction of interstitial fluid and 5% blood vessels (primarily capillaries) and is therefore be a credible source of glucose concentration. Of course, one should be able to obtain sufficient photons in the transmission geometry in order to exploit the aforementioned feature.

Indeed, transmission Raman spectroscopy can be perceived as a special case of the recently proposed spatially offset Raman spectroscopy. This method relies solely on the spatial properties of the (diffuse component) emerging Raman photons at the sample surface, similar to tomographic concepts used extensively for NIR absorption [9] and fluorescence spectroscopy [10]. It is essentially based on the collection of Raman spectra at spots that are spatially offset from the excitation spot on the surface of the sample. Evidently, the photons traverse to deeper distances are also more likely to diffuse sideways than the Raman photons that originate close to the surface. Using this principle, Matousek and co-workers have published a series of papers demonstrating the capability of SORS to provide better SNR for analytes embedded deeper in the turbid medium while rejecting most of the signal emanating close to the surface [11-15]. SORS, which can be performed using the same fiber probe geometry of central excitation and annular ring of collection fibers in the reflection geometry, provides better signals than the analogous time gated approach because of the restrictive signal acquisition capability of the latter.

Additionally, SORS also enables the suppression of surface fluorescence (e.g. fluorescence arising from melanin in skin tissue) in a similar fashion to that of surface Raman.

Adopting such an approach may produce substantive benefits for blood analyte monitoring as well. Ultimately, we envision that the site specific turbidity as well as its anatomy will play a crucial role in determining the parameters employed for SORS (especially the excitation collection separation) and whether or not transmission Raman would be a suitable modality.

3. Design and fabrication of a suitable sample interface

A crucial feature of the overall instrumentation, which has hitherto not received careful consideration, is the design of the sample interface. This represents an engineering problem with several functional requirements, all of whom must be carefully considered before finalizing the design parameters. We outline the major factors that should be considered in order to develop a sample interface for a fiber-probe based system.

First, the question of whether to proceed with a contact probe or a non-contact (standoff) probe must be resolved. The former approach would introduce questions of contact pressure and temperature as well as the refractive index change between the probe tip and the skin surface. The non-contact mechanism would not have issues of contact pressure and temperature but would have to undergo two sets of refractive index changes, namely at the probe/air and air/skin tissue interfaces. In this context, optical clearing (with glycerol, for example) could be pursued in order to obtain greater penetration depth into the sample [17], although signals emanating from the clearing agent could add to the complexity of the intrinsic Raman spectrum.

Second, for a person specific instrument (which represents our ultimate goal in developing a miniaturized instrument) several of the confounding factors including tissue turbidity and

heterogeneity (e.g. differences in density and geometry of blood vessels in the dermis) can be substantially reduced by enabling measurements at the same site location. Ideally, we envision a band around the wrist or a skin surface patch where the fiber probe could be attached periodically to perform the non-invasive blood glucose measurements. While any such band or patch is likely to undergo a finite level of physical displacement (at least on the micron scale), the reproducibility of spectral measurement would be greatly enhanced by having a largely overlapping sampling volume. It is worth noting that the long term stability of such a physical interface (subject to sweating and other physiological changes) would have to carefully studied.

Third, an intriguing approach for signal enhancement of the blood analytes is to employ an active mechanism such as tissue modulation [17, 18]. As stated by Chaiken and co-workers [19], "tissue modulation is the use of spatiotemporally localized mechanical, thermal, chemical and/or other external influences to manipulate the mobile components of tissue relative to the static components." The fundamental concept is to use difference spectroscopy to isolate the spectra of the mobile and static tissues. More specifically, tissue modulation can be utilized to obtain chemical information concerning subsurface tissues as the surface tissues are static (they deform but do not move, per se). Since the primary mobile component is the blood and the analytes present in the blood, it has been reported that this approach would afford greater specificity in terms of selecting the signal of the analyte of interest while rejecting the baseline features of the other skin tissue components. Evidently, the proposed approach is likely to exhibit maximum effectiveness in applications where the baseline signal is much greater than the signal from the analyte of interest, as is the case for blood glucose monitoring. Also, in order for this method to work, the active (or external) fluctuations should be larger than the natural fluctuations stemming from blood flow and skin tremor. In this context, one could design a mechanical interface that

adjusts the contact pressure to allow for different quantities of blood volume to pool in a specific site and then use difference spectroscopy before further chemometric analysis.

4. Incorporation of additional non-invasive modalities

Despite the specificity of Raman spectroscopy in detecting various bioanalytes, its sensitivity has hindered its widespread application in the real-world. And even though our uncertainty estimates provide confidence in being able to measure clinically relevant concentrations accurately and precisely, one way to complement the proposed diagnostic method would be to incorporate an additional optical/spectroscopic scheme employing identical excitation-collection geometry.

The natural choice would be NIR absorption spectroscopy, because of it provides complementary vibronic signatures to that of Raman. Further, its instrumentation is relatively straightforward and can be readily incorporated with the existing Raman unit. (It should be mentioned that addition of any modality would conflict with the notion of miniaturization but may be a useful step in demonstrating clinical feasibility of a non-invasive system.) Moreover, there exists a rich history of the application of NIR absorption spectroscopy for non-invasive blood glucose detection (Chapter 2), largely because of the much higher signals acquired. While NIR absorption spectroscopy lacks the specificity afforded by Raman (as discussed earlier in the thesis), the combination of NIR absorption (high sensitivity, low specificity) and Raman (low sensitivity, high specificity) may provide an ideal platform for non-invasive glucose monitoring.

Clinical directions

1. Spectroscopic Glucose Tolerance Test (SGTT)

The oral glucose tolerance test (OGTT), as has been described elsewhere in this thesis, is one of the most widely used clinical screening procedures for the detection of diabetes and related

conditions. Typically, blood is drawn at short intervals over a 2-3 hour period to measure the subject's insulin response to an oral glucose challenge. Over the time course of the measurements, healthy volunteers will show physiological recovery to euglycemic levels due to insulin response, while recovery is significantly delayed or is absent in individuals with various forms of diabetes.

We believe that Raman spectroscopy can non-invasively provide the required kinetic information regarding insulin response and replace repeated blood withdrawals in OGTT. Initially, one can demonstrate the feasibility of spectroscopic glucose tolerance test (SGTT) by simultaneous Raman spectroscopy and blood withdrawal in 5 healthy volunteers and 5 diabetic patients. If the results show a clinically acceptable correspondence, i.e. +/- 20% variation between direct blood glucose measurements and Raman spectroscopic predictions (ISO Standard 15197) [20], we will prospectively demonstrate the clinical applicability of SGTT by a blinded comparison between SGTT and OGTT. Importantly, since the kinetic profile of glucose concentrations - rather than their absolute values - provides sufficient diagnostic information during an OGTT, this application is likely to represent the first clinical implementation of this new technology. The exact number of patients in the blinded comparison can be determined using standard power calculations once the initial performance parameters for spectral-based glucose measurements are determined.

2. Clinical measurements of other blood analytes

Future research could also focus on the non-invasive measurements of other clinically important blood analytes, especially in a hospital setting such as in a triage or intensive care unit (ICU). Specifically, creatinine and urea measurements would provide an assessment of kidney function and could be readily detected due to their characteristic Raman signatures. Levels of creatinine,

which is a break-down product of creatine phosphate in muscle, are typically employed for computing the creatinine clearance, which reflects the glomerular filtration rate (GFR). The GFR is clinically important because it is a measurement of the ability of the kidney to filter blood and produce urine.

Furthermore, in order to obtain a more complete understanding of the renal function, the blood concentrations of both creatinine and urea is necessary. Additionally, the blood urea nitrogen-to-creatinine ratio (urea concentrations are typically expressed in terms of blood urea nitrogen, BUN) can indicate other problems besides those intrinsic to the kidney. A pre-renal problem, e.g. volume depletion, where the urea level is raised disproportionately with respect to the creatinine concentration provides one such example. While continuous measurements of these analytes are not required in general (in contrast to glucose measurements in diabetic patients), they represent an important class of diagnostic problems in the ICU and in neonates, where frequent blood withdrawal is either not possible or may potentially induce serious complications.

Similarly, one could extend the scope of study to perform non-invasive alcohol and drug testing, which is an important screening procedure in the triage. Given the larger Raman scattering cross-sections of alcohol and a large number of drugs, our current instrumentation and methodologies are likely to be sufficient for a proof-of-concept demonstration in point-of-care studies in human subjects overdosed on alcohol or/and on specific drugs.

9.3 References

(1) Haka, A. S.; Volynskaya, Z.; Gardecki, J. A.; Nazemi, J.; Shenk, R.; Wang, N.; Dasari, R. R.; Fitzmaurice, M.; Feld, M. S. *Journal Biomedical Optics* **2009**, 14, 054023.

(2) Volynskaya, Z.; Haka, A. S.; Bechtel, K. L.; Fitzmaurice, M.; Shenk, R.; Wang, N.; Nazemi, J.; Dasari, R. R.; Feld, M. S. *Journal of Biomedical Optics* **2008**, 13, 024012.

(3) Scepanovic, O. R.; Volynskaya, Z.; Kong, C. R.; Galindo, L. H.; Dasari, R. R.; Feld, M. S. *Review of Scientific Instruments* **2009**, 80, 043103.

(4) Utzinger, U.; Richards-Kortum, R. R. *Journal of Biomedical Optics* **2003**, 8, 121.

(5) Motz, J. T. *Development of in vivo Raman spectroscopy of atherosclerosis*, Ph.D. Thesis, Massachusetts Institute of Technology, **2003**.

(6) Winston, R.; Minano, J. C.; Benitez, P. G. *Nonimaging Optics*, Elsevier Academic Press, Burlington, MA, **2005**.

(7) Tanaka, K.; Pacheco, M.; Brennan, J.; Itzkan, I.; Berger, A. J.; Dasari, R.; Feld, M. *Applied Optics* **1996**, 35, 758-763.

(8) Burmeister, J. J.; Arnold, M. A. *Clinical Chemistry* **1999,** 45, 1621 - 1627.

(9) Das, B. B.; Liu, F.; Alfano, R. R. *Reports on Progress in Physics* **1997**, 60, 227–292.

(10) Pfefer, T. J.; Schomacker, K. T.; Ediger, M. N.; Nishioka, N. S. *Applied Optics* **2002**, 41, 4712–4721.

(11) Matousek, P. *Chemical Society Reviews* **2007**, 36, 1292–1304.

(12) Matousek, P.; Clark, I. P.; Draper, E. R. C.; Morris, M. D.; Goodship, A. E.; Everall, N.; Towrie, M.; Finney, W. F.; Parker, A. W. *Applied Spectroscopy* **2005**, 59, 393–400.

(13) Matousek, P.; Morris, M. D.; Everall, N.; Clark, I. P.; Towrie, M.; Draper, E.; Goodship, A.; Parker, A. W. *Applied Spectroscopy* **2005**, 59, 1485–1492.

(14) Schulmerich, M. V.; Finney, W. F.; Fredricks, R. A.; Morris, M. D. *Applied Spectroscopy* **2006**, 60, 109–114.

(15) Matousek, P.; Draper, E. R. C.; Goodship, A. E.; Clark, I. P.; Ronayne, K. L.; Parker, A. W. *Applied Spectroscopy* **2006**, 60, 758–763.

(16) Plotnikov, S; Juneja, V; Isaacson, A. B.; Mohler, W. A.; Campagnola, P. J. *Biophysical Journal* **2006**, 90, 328-339.

(17) Wood, E. *U.S. Patent No. 2,706,927 26*, **1955**.

(18) Chaiken, J.; Finney, W.; Knudson, P. E.; Weinstock, R. S.; Khan, M.; Bussjager, R. J.; Hagrman, D.; Hagrman, P.; Zhao, Y. W.; Peterson, C. M.; Peterson, K. *Journal of Biomedical Optics*, **2005**, 10, 031111.

(19) Chaiken, J.; Rudd, K.; Voss, E.; Bussjager, R. J.; Rice, D.; Godici, D. *US Patent Application No. 20080306363A1*, **2008**.

(20) Bohme, P.; Floriot, M.; Sirveaux, M.; Durain, D.; Ziefler, O.; Drouin, P.; Guerci, B. *Diabetes Care*, **2003**, 26 (4), 6.